Asymmetrical Learning of Win and Loss Associations:
Individual Differences and Task Effects

by

Lilian Edith Cabrera-Haro

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Psychology)
in the University of Michigan
2021

Doctoral Committee:

Professor Patricia A. Reuter-Lorenz, Chair
Professor John E. Laird
Professor Taraz Lee
Professor Richard L. Lewis

Lilian E. Cabrera-Haro

lilianec@umich.edu

ORCID iD: 0000-0001-7478-538X

Dedication

To my daughter, husband, parents, and sisters.

Thank you for your love, never-ending encouragement, and unwavering support.

Acknowledgements

First, I would like to extend my deepest gratitude to my advisor, Patricia Reuter-Lorenz for her guidance and support throughout my time at the University of Michigan. I am extremely thankful for all of the knowledge she has shared with me, the opportunities that she has afforded me, and her dedication to my scientific training. It has been an honor and incredible privilege to work under such an extraordinary scientist and individual.

Next, I would like to thank my dissertation committee members, Richard Lewis, Taraz Lee, and John Laird for their feedback and valuable insight throughout the dissertation process.

I would also like to thank my fellow CCN graduate student family. In particular I would like to thank Colleen Frank, Tessa Abagis, Myrna Cintron-Valentin, Tiffany Jantz-Desormeaux, Jordan Meyer, Ziyong Lin, and Stella Hao. They have been a tremendous source of support, knowledge, and friendship.

Finally, on a more personal note, I would like to thank my parents who gave me a strong foundation so that I could pursue a higher education. My husband, Joshua Haro, for providing his unwavering support and encouragement when I needed it the most. To my daughter, Juliette, who has given me the strength and drive to push through challenges. Without your love and support, none of this would be possible. Thank you.

Table of Contents

Dedication ii

Acknowledgements iii

List of Tables vi

List of Figures viii

Abstract xi

Chapter 1 Introduction 1

Value-Driven Attention 2

Learning Asymmetry 4

Individual Differences in Learning Win and Loss Associations 6

Conditions Leading to Valence-Based Asymmetries 7

Overview of the Present Dissertation 8

Chapter 2 Intraindividual Consistency and Individual Differences in Learning of Win and Loss Associations 11

Introduction 11

Method 19

Results 25

Discussion 41

References 49

Supplemental Material 53

Chapter 3 Separating Learning for Win and Loss Associations 55

Introduction 55

Method 59

Results 64

Discussion 77

References 85

Supplemental Material 88

Chapter 4 The Effect of Outcome Magnitude on Win and Loss Associations 100

Introduction 100

Experiment 1 104

Method 104

Results and Discussion 108

Experiment 2 115

Method 115

Results 121

General Discussion 134

Supplemental Material 145

Chapter 5 General Discussion 154

Summary of Key Findings 155

Implications for Boundary Conditions on the Learning Asymmetry 161

Implications for Research on Individual Differences in Win and Loss Learning 163

Implications for Theories of Decision-Making 164

Limitations and Future Directions 166

Closing Remarks 169

References 171

List of Tables

Table 2.1. Pearson r Values for the Correlations between Questionnaire Measures and VLT Performance. 31

Table 2.2. Pearson r Values for the Correlations between Questionnaire Measures and RL Parameters. 32

Table 2.3. Multinomial Logistic Regression Predicting Outcome Response for each VLT scene from Valence and Probability (N = 56) 33

Table 3.1. Binary Logistic Regression Predicting Memory Accuracy from Valence, Optimality, and Valence Order. 70

Table 3.2. Multinomial Logistic Regression Predicting Outcome Response for each VLT scene from Valence, Probability, and Valence Order (N = 120) 75

Table 3.3. Learning Rate and Inverse Temperature Parameter for each Valence Order Condition.. 77

Table 3.4. Binary Logistic Regression Predicting Memory Accuracy from Valence, Optimality, and Valence Order. 92

Table 3.5. Multinomial Logistic Regression Predicting Outcome Response for each VLT scene from Valence, Probability, and Valence Order (N = 196) 97

Table 3.6. Learning Rate and Inverse Temperature Parameter for each Valence Order Condition.. 99

Table 4.1. Binary Logistic Regression Predicting Memory Accuracy from Valence and Optimality. 111

Table 4.2. Multinomial Logistic Regression Predicting Outcome Response for each VLT scene from Valence and Probability for Learners. 113

Table 4.3. Pearson r values for the Correlation between the Learning Difference Score and Measures of Sensitivity to Punishment and Reward and Loss Aversion. 125

Table 4.4. Binary Logistic Regression Predicting Memory Accuracy from Valence, Optimality, and Outcome Magnitude. 126

Table 4.5. Means and Standard Deviation of Memory Accuracy for the VLT scenes. 126

Table 4.6. Multinomial Logistic Regression Predicting Outcome Response for each VLT scene from Valence and Probability. 128

Table 4.7. Binary Logistic Regression Predicting Accuracy of Outcome Magnitude from Valence, Optimality, and Outcome Magnitude. 130

Table 4.8. Pearson r Values for the Correlations between Questionnaire Measures and RL Parameters. 134

Table 4.9. Correlational Analyses between the Learning Difference Score and Measures of Sensitivity to Punishment and Reward and Loss Aversion. 147

Table 4.10. Binary Logistic Regression Predicting Memory Accuracy from Valence, Optimality, and Outcome Magnitude. 148

Table 4.11. Multinomial Logistic Regression Predicting Outcome Response for each VLT scene from Valence and Probability. 150

Table 4.12. Binary Logistic Regression Predicting Accuracy of Outcome Magnitude from Valence, Optimality, and Outcome Magnitude. 152

List of Figures

Figure 2.1. Study protocol for each session. ... 20

Figure 2.2. Schematic of the Value Learning Task. ... 23

Figure 2.3. Schematic of the Memory Task. ... 25

Figure 2.4. Value Learning Task Performance. ... 27

Figure 2.6. Scatterplot for the Learning Difference Score between Session 1 and 2. ... 29

Figure 2.5. Learning difference scores for each participant, sorted by the learning difference score during Session 1. ... 29

Figure 2.7. Scatterplot for win learning between Session 1 and Session 2. ... 30

Figure 2.8. Scatterplot of loss learning between Session 1 and 2. ... 30

Figure 2.9. Memory Task Performance. ... 33

Figure 2.10. Outcome Attributions for Win Scenes. ... 34

Figure 2.11. Outcome Attributions for Loss Scenes. ... 35

Figure 2.12. Learning difference score and reinforcement learning parameters for each participant during Session 1. ... 37

Figure 2.13. Learning difference score and reinforcement learning parameters for each participant during Session 2. ... 38

Figure 2.14. Scatterplots of model-derived parameters and the learning difference score. ... 39

Figure 2.16. Scatterplots of the inverse temperature parameter across sessions. ... 40

Figure 2.15. Scatterplot of the learning rate across sessions. ... 40

Figure 2.17. Value Learning Task Performance for Lin et al. (2020) data set. ... 54

Figure 3.1. Sample trial of the Value Learning Task. ... 62

Figure 3.2. Schematic of the Memory Task. ... 64

Figure 3.3. Probability of optimal choice in the Value Learning Task based on Valence Order. 67

Figure 3.4. Memory data accuracy as a function of Valence and Probability of Outcomes 70

Figure 3.5. Outcome attribution responses on the Memory Task for win scenes........................ 75

Figure 3.6. Outcome attribution responses on the Memory Task for loss scenes. 76

Figure 3.7. Probability of optimal choice for each Valence Order condition collapsed across valence. 89

Figure 3.8. Memory data accuracy as a function of Valence and Probability of Outcome across all valence order conditions. 92

Figure 3.9. Outcome attribution responses on the Memory Task for win scenes........................ 97

Figure 3.10. Outcome attribution responses on the Memory Task for loss scenes. 98

Figure 4.1. Schematic of the Value Learning Task. 106

Figure 4.2. Schematic of the Memory Task.. 108

Figure 4.3. Performance on the Value Learning Task for participants who reached the learning criterion. 110

Figure 4.4. Accuracy of explicit knowledge of outcomes associated with the VLT scenes....... 112

Figure 4.5. Outcome attributions for the win scenes. 114

Figure 4.6. Outcome attributions for the loss scenes. 114

Figure 4.7. Schematic of the Value Learning Task in the Outcome Magnitude Experiment..... 118

Figure 4.8. Schematic of the Memory Task in the Outcome Magnitude Experiment. 120

Figure 4.9. Interaction between Outcome Magnitude and Trial Block. 123

Figure 4.10. Probability of optimal choice was significantly higher for the win pair versus loss pair in each outcome magnitude condition. 123

Figure 4.11. Performance on the VLT for participants who met the learning criterion. 124

Figure 4.12. Accuracy of explicit knowledge of outcomes associated with the VLT scenes..... 127

Figure 4.13. Outcome attributions in the Memory Task... 129

Figure 4.14. Outcome magnitude accuracy. 131

Figure 4.15. Scatterplots of learning difference score and reinforcement learning parameters. 133

Figure 4.16. Performance on the VLT for participants who did not meet the learning criterion..... 146

Figure 4.17. Accuracy of explicit knowledge of outcomes associated with the VLT scenes..... 148

Figure 4.18. Outcome attributions in the Memory Task.. 151

Figure 4.19. Outcome magnitude accuracy. 152

Figure 4.20. Scatterplots of learning difference score and reinforcement learning parameters. 153

Abstract

Learning that particular objects or actions are associated with rewards and punishment (i.e., value) can impact many aspects of human behavior, including what we attend to, what we desire, and what we learn. A recent report (Lin et al. 2020) documented an asymmetry in the learning of win versus loss associations in a widely used value learning task (VLT; Raymond and O'Brien, 2009). In this task people learn, via trial and error, to associate otherwise neutral stimuli with win or loss outcomes of varying probability. However, a number of questions about the potential boundary conditions of this learning asymmetry remained unanswered. The goal of the present dissertation was to investigate whether better learning of win associations than loss associations can be attributed, in part, to individual differences that are consistent over time, whether the asymmetry depends on the context in which people learn win and loss associations, and whether the asymmetry depends on the magnitude of the win and loss outcomes. To this end, participants completed the VLT (1) on two separate occasions, (2) when win, loss, and no change trials were presented in separate trial blocks, and (3) when the magnitude of the win/loss outcomes varied. The results showed that while individuals differed in the extent to which they learned win associations better than loss associations, there was no strong evidence for intraindividual consistency in learning patterns or that learning patterns correlated with other individual difference variables known to affect decision behavior. However, evidence that learning context influenced the win-loss asymmetry emerged across all experiments. First, when participants completed the VLT in two sessions, the learning asymmetry was reduced during the second session. Second, when separating win and loss trials, we found that the learning asymmetry only

emerged when loss trials directly preceded win trials. Third, we demonstrated that including multiple outcome magnitude conditions in the same context affects the learning pattern for win and loss associations. Taken together these findings reveal that while the learning asymmetry is robust, there are conditions that improve loss learning and reduce the asymmetry. Thus, loss learning is particularly malleable to the context in which value learning takes place. When examining explicit knowledge of the learned associations, we found that whether or not a learning asymmetry emerged in the VLT, explicit knowledge of the outcomes differed for win and loss scenes. We suggest that the memory representation of stimulus-outcome associations may depend, at least in part, on choice history. Lastly, when examining the association between reinforcement learning parameters and the learning asymmetry, we found that a higher learning rate was associated with a smaller asymmetry. Whereas the results for the inverse temperature (balance between exploration vs. exploitation) were mixed. We suggest that the conditions that affect the presence or absence of the learning asymmetry may also influence whether exploring versus exploiting is beneficial for learning the optimal choice. Future work can examine whether selectively increasing the loss value would equate wins and losses, thus reducing the asymmetry. Furthermore, future studies examining subsequent cognitive consequences of acquired win and loss associations of otherwise neutral stimuli should consider the learning context and how choice history plays a role in subsequent effects of learned value.

Chapter 1

Introduction

Our environment places demands on our attentional system that exceed our capacity. Imagine a scenario where you go the airport to pick up your friend. While the airport is busy with hundreds of guests arriving and departing, you manage to spot your friend among a crowd of individuals. How is that you were able to direct your attention to find your friend despite a backdrop of an overwhelming amount of visual information? One mechanism that drives how attention is allocated is the ‘value’ associated with that information (Anderson, 2016). Value can be defined subjectively, such as by the personal relevance of the information. Or, value can be defined objectively based on a quantitative metric, such as the number of points or amount of money information is associated with. To study how value affects how we perceive, attend, and act on information, value can be experimentally manipulated in the laboratory by establishing an association between an otherwise neutral stimulus and an outcome that varies in valence (e.g., rewards versus losses), the probability that a certain outcome will occur (e.g., low versus high probability of earning a reward), and the magnitude of the outcome (e.g., low versus high reward).

The present dissertation investigates value learning, which refers to the process by which associations are formed between an action and outcome in order to maximize rewards and minimize losses (Knutson et al., 2001; O’Doherty, Deichman, Critchley, & Dolan, 2002; Raymond & O’Brien, 2009). More specifically, the set of experiments reported in the subsequent chapters examine value learning patterns for stimuli that lead to wins and losses, and the conditions under which a learning difference emerges for win and loss associations.

Value-Driven Attention

In the attention domain, researchers have focused on understanding how we manage our limited information processing ability when we are faced with a large amount of sensory input. Theories about how we manage our information processing limitations have distinguished between selective attention that is goal-directed (top-down) or stimulus-driven (bottom-up). According to such theories, attentional selection results from an interplay of voluntary, goal-directed processes, driven by the alignment of information with context-specific goals, and automatic, stimulus-driven processes driven by the intrinsic, physical properties of stimuli (e.g., color, brightness, loudness) (Corbetta & Shulman, 2002; Posner, 1980; Yantis, 2000; Yantis & Jonides, 1984). While early theories of attention have focused on this dichotomy (e.g., Corbetta & Shulman, 2002; Theeuwes, 2010) more recent research has examined how the value of information, acquired through stimulus-outcome associations, can also drive what information we attend to or ignore (Anderson, Laurent, & Yantis, 2011; Della Libera & Chelazzi, 2006; 2009; Raymond & O'Brien, 2009). The benefit of learning the value of a stimulus is that people can selectively attend to information so as to maximize positive outcomes and avoid negative outcomes. Therefore, a more recent body of research has been dedicated to understanding and characterizing value-driven attention.

Much of the early work focused on examining how positive outcomes (i.e., rewards) influence selective attention. Using a priming task, Della Libera and Chelazzi (2006) demonstrated that participants' ability to select a target was impaired if they were highly rewarded for ignoring that stimulus on a previous trial. More compelling evidence of the influence of rewards on attention came from a study demonstrating value as a unique mechanism for driving selective attention (Anderson, Laurent, & Yantis, 2011). In their study, participants first learned associations between color targets and a monetary reward. In a subsequent phase, the same color stimuli served as distractors during search for a shape-defined target. The previously reward-associated distractors were neither physically salient nor task relevant, therefore ruling out a stimulus-driven or goal-directed attentional mechanism. The results demonstrated attentional capture by the previously reward-associated distractors, providing evidence for a unique value-driven attentional selection mechanism. Since then, numerous studies have been conducted further characterizing value-driven attention, including how stimuli associated with punishment may also modulate selective attention.

One approach to studying value-driven attention is to examine how attention is allocated toward stimuli previously associated with rewards or punishment, after the reward schedule is discontinued and no longer task relevant. These studies typically split the experiment in two phases. During the first phase, participants learn via trial and error, that otherwise neutral stimuli are associated with wins and losses of varying probability. During the second phase, participants complete a second unrelated task where the stimuli from the first phase appear, but the reward schedule is discontinued. Researchers use the second phase to examine how valence (wins versus losses) and predictiveness (low versus high probability that an outcome will occur) affects different cognitive processes, such as perceptual processing (O'Brien & Raymond, 2012; Rutherford, O'Brien & Raymond, 2010), attention (Raymond & O'Brien, 2009) and motor control (e.g., Painter, Kritikos, & Raymond, 2013).

While differential effects of acquired value on subsequent stimulus processing have been reported, the question of whether initial learning of win and loss associations is equivalent has been relatively unexplored. However, a recent investigation reported a robust learning difference between win and loss associations, such that win associations were learned better than loss associations (Lin, Cabrera-Haro, & Reuter-Lorenz, 2020). This learning asymmetry poses a potential challenge for the conclusions drawn about the effects of learned value on subsequent stimulus processing. If learning for win and loss associations is not equivalent, then any learning differences, rather than acquired value, could be the basis for valence effects that arise in subsequent tasks.

Learning Asymmetry

The robustness of the learning asymmetry for win and loss associations was recently reported in a meta-analysis of studies using a probabilistic learning task developed by Raymond and O'Brien (2009), or similar value learning task (Lin et al., 2020). Furthermore, in two new empirical studies, the authors demonstrated that this learning asymmetry was evident when wins and losses led to either point or monetary earnings, and whether or not participants received explicit instructions about the outcome contingencies. Altogether, the results of the meta-analysis and empirical study from Lin et al. 2020 suggest that although these probabilistic learning tasks incorporate symmetrical outcome probabilities for wins and losses and equivalent number of learning trials for each condition, equivalent learning for win and loss associations cannot be assumed. However, the reason(s) for the learning asymmetry remained unclear.

To investigate the nature of the learning asymmetry, a reinforcement learning model was implemented to simulate participants' learning and performance on the value learning task using the data from Lin et al. (2020) (Hao, Cabrera, Lin, Reuter-Lorenz, & Lewis, 2019). A computational approach allows us to characterize different learning styles for win and loss associations with parameters such as the learning rate and degree to which participants explore versus exploit what they have learned on a trial-by-trial basis (Sutton & Barto, 1998). The value learning task is structured such that the win and loss pairs have equivalent absolute expected value (reward x probability of earning the reward).[1] Learning the value of each stimulus involves trial and error in selecting a stimulus and receiving feedback. With repeated encounters (i.e., selecting a stimulus within each pair), the frequency with which an outcome (win, loss, or no change) follows selection leads to the stimulus acquiring value. When performance on the value learning task was simulated on a trial-level basis, the results from the reinforcement model yielded the same learning asymmetry observed by Lin et al. (2020) despite equivalent absolute expected value for the win and loss pairs. This learning asymmetry was attributed to value estimates that were more poorly discriminated for stimuli in the loss pair than stimuli in the win pair. Even when performance was simulated with the optimal parameters corresponding to an aggressive learning rate and balance between exploring and exploiting what is learned, the asymmetry persisted, though was diminished considerably. While this modeling work provides insight into the basis for the learning asymmetry, a number of questions about the boundary conditions of this learning asymmetry remain unanswered. The work described in this dissertation aims to explore (1) whether there are underlying individual differences in the extent to which win associations are learned better than loss associations, (2) whether learning patterns for wins and losses are consistent across sessions, (3) whether the context in which win and loss trials appear affects learning, and (4) whether the learning asymmetry is reduced when the magnitude of win and loss outcomes increase.

[1] Note that in the context of reinforcement learning, reward refers to both positive and negative outcomes.

Individual Differences in Learning Win and Loss Associations

The observation that learned value can bias information-processing has motivated research to examine whether certain individuals may be more prone to value-driven attentional biases. For example, patients with previous or current drug-dependencies showed distraction by stimuli previously associated with nondrug (monetary) reward (Anderson et al. 2013; 2016). On the other hand, value-driven attentional capture has been shown to be blunted in individuals with depressive symptoms (Anderson et al. 2014).

Another area of research in individual differences has focused on investigating whether people differ in their tendencies to seek positive outcomes or to avoid negative ones. This research has demonstrated that whether participants approach positive outcomes more than avoid negative ones (or vice versa), is associated with factors such as particular genetic polymorphisms (e.g., Frank & Hutchison, 2009), different levels of dopamine function (e.g., Frank, Seeberger, & O'Reilly, 2004), and different approach and avoidance personality styles (e.g. Aberg, Doell, & Schwartz, 2016).

Regarding individual differences in learning win and loss associations in the value learning task, exploratory analyses of learning differences using the data in Lin et al. (2020), revealed a subgroup of individuals who learned win and loss associations equally well and individuals who learned losses better than wins. This finding suggests that the advantage for learning win relative to loss associations is not obligatory. Therefore, this exploratory analysis raises the question of whether the learning asymmetry is universal, and a potential property of the VLT, or whether there are individual difference variables that influence this pattern. Given that previous research has demonstrated that people differ in their tendencies to seek positive outcomes and avoid negative outcomes, it is possible that there are underlying individual differences in the extent to which people learn win associations better than loss associations. Therefore, the first aim of the present dissertation was to test the hypothesis that there are underlying individual differences in the extent to which gains are learned better than losses. This hypothesis is tested in Chapter 2 of this dissertation.

Conditions Leading to Valence-Based Asymmetries

In addition to examining individual differences in learning about positive versus negative outcomes, a second question concerns whether the conditions under which learning for win and loss associations occurs affects the emergence of a learning asymmetry. A breadth of research examining other valence-based asymmetries have explored conditions under which behavioral responses to positive and negative outcomes differ. For example, prior research has demonstrated that whether gains and losses appear in the same condition versus separate conditions can modulate how people psychologically or physiologically respond to losses versus gains (for a review see Yechiam & Hochman, 2013). While these asymmetric effects have been explored primarily in decision-making research involving risk-taking behavior of gains and losses, whether these effects also operate in probabilistic learning tasks has been unexplored. In the standard value learning task (Raymond & O'Brien, 2009), trials with win, loss, and no change outcomes are randomly interleaved requiring that people learn all associations concurrently. While the learning advantage for wins compared to losses has appeared when trials with win and loss outcomes are intermixed, this learning asymmetry may be altered when learning for wins and losses is blocked. This possibility is examined in Chapter 3 of this dissertation.

Another condition that may influence the asymmetry is the magnitude of wins and losses. One robust observation of how people respond differently to gains versus losses is that people prefer to avoid losses compared to acquiring gains of equal value (e.g., +$5.00 versus $-5.00) suggesting that losses loom larger than gains (*loss aversion*; Kahneman & Tversky, 1979). Although this finding has received empirical support in a variety of decision-making research domains (e.g., economic and political decision-making), recent work has identified conditions under which loss aversion is less prominent. For example, loss aversion can be reduced or reversed when the magnitude of outcomes is small (e.g., Harinck, Van Dijk, Van Beest, & Mersmann, 2007). One potential explanation for such reversals is that because small losses are more common than large losses, it is easier to trivialize small losses compared to large losses (e.g., compare the feeling of losing $0.20 to losing $200) (Stewart, Chater, & Brown, 2006). On the other hand, gains are readily accepted even when small in magnitude. Consequently, for small amounts of money, the positive feelings associated with gains outweigh the negative feelings associated with losses. In other words, gains loom larger than losses. Whereas, for large amounts of money, the positive feelings associated with large positive outcomes do not outweigh the negative feelings associated with large losses. In other words, losses loom larger than gain.

In the studies adopting the VLT, or similar probabilistic value learning tasks, that were included in the meta-analysis (Lin et al., 2020), the magnitude of wins and losses were of relatively small magnitude (e.g., +/- $ 0.50). To our knowledge, the effect of outcome magnitude on learning win and loss associations has not been tested in the VLT. Therefore, we hypothesize that increasing the magnitude of win and loss outcomes may reduce or reverse the observation that win associations are learned better than loss associations. This hypothesis is tested in Chapter 4.

Overview of the Present Dissertation

Many of the decisions that we make on a daily basis involve balancing factors that will lead to desirable outcomes while avoiding behaviors with potential negative outcomes. These decisions are informed from past experience of outcomes that result from specific actions. Therefore, learning that certain information is associated with rewards or punishments is beneficial to guide future behavior in actions that we may take. The focus of this dissertation is to further understand people's learning for win and loss outcome associations.

The first aim of this dissertation is to investigate whether the learning asymmetry that was previously reported by Lin et al. (2020) is a dimension along which individuals vary and whether learning patterns will be consistent within individuals. The experiment reported in Chapter 2 was designed to test the hypothesis that the learning asymmetry for win and loss associations can be attributed to individual differences in learning for win and loss associations by having participants complete the value learning task in two separate occasions. If the learning asymmetry can be attributed to stable individual differences, then learning for win and loss associations should be consistent across the two sessions.

The second aim of this dissertation is to investigate whether the learning asymmetry depends on the context in which win and loss associations are learned. The experiment reported in Chapter 3 was designed to test the hypothesis that the learning asymmetry depends on the context in which wins and losses are presented. Win and loss associations were presented in separate trial blocks to test whether separating learning may reduce the learning asymmetry.

The third aim of this dissertation is to investigate whether the magnitude of outcomes affects the learning pattern for win and loss associations. The experiment reported in Chapter 4 was designed to investigate whether the learning asymmetry depends on the magnitude (low versus high) of win and loss outcomes. The VLT was modified to include trials where wins and losses led to low magnitude outcomes (5 points) and high magnitude outcomes (50 points). The findings from the three experiments are summarized and their theoretical implications are discussed in Chapter 5.

This body of work will provide evidence that the learning asymmetry is remarkably robust, persisting across multiple sessions, whether win, loss, and no change pairs are presented in different trial blocks, and with low or high magnitude outcomes. While there was little evidence for intraindividual consistency in learning patterns, there were conditions where the learning asymmetry was reduced or eliminated.

Chapter 2

Intra-individual Consistency and Individual Differences in Learning of Win and Loss Associations

Introduction

Human behavior is in part guided by the motivation to maximize rewarding outcomes and minimizing unpleasant states or punishment (Daw et al., 2006). However, people vary considerably in the degree to which their behavior is motivated by reward-relevant stimuli and inhibited by punishment-relevant stimuli (e.g., Frank, Woroch, & Curran, 2005; Scheres & Sanfey, 2006). Individual differences are also apparent in neural activation patterns associated with positive and negative outcomes. For example, when examining neural responsiveness to reinforcement of gains and losses, research has shown that, in general, activation in the left and right ventral striatum is associated with reward reception. In contrast, activation of the dorsal striatum is associated with learning to avoid punishment in a reinforcement learning task (Kim et al. 2015). However, these activation patterns can vary based on individual differences in sensitivity to reward and punishment. Kim et al. (2015) examined how trait sensitivity to reward and punishment, as measured by a questionnaire, was correlated with activation of neural regions while participants were completing a reinforcement learning task. Greater trait sensitivity to reward was associated with greater activation in the left and right ventral striatum during reward reception. In contrast, greater trait sensitivity to punishment was associated with less activation in the left dorsal striatum during avoidance anticipation. These results suggest that individuals who differ in trait sensitivity to reward and punishment show different activation patterns of neural regions associated with learning to attain gains and avoid losses.

Researchers have also explored individual differences in learning from reward and punishment at the behavioral level. For example, using a reinforcement learning paradigm in which participants learned the association between cues and probabilistic outcomes, researchers have documented age differences in how people learn from reward versus punishment (Palminteri et al., 2016). Adults learned symmetrically from both reward and punishment, whereas adolescents learned from reward but were less likely to learn from punishment. Learning differences have also been attributed to psychiatric conditions. In a probabilistic classification task involving positive and negative feedback, patients with Social Anxiety Disorder and General Anxiety Disorder learned better from negative feedback than either patients with Panic Anxiety Disorder or healthy control participants (Khdour et al. 2016). Other individual differences in learning from reward and punishment have been attributed to varying levels of cognitive impulsivity (e.g., Cáceres & San Martín, 2018), and trait approach and avoidance behaviors (Aberg et al. 2016). Altogether, these findings suggest that people can vary in the extent to which they respond to reward versus negative outcomes.

Intraindividual Consistency in Value Learning

A recent investigation of how people learn to associate neutral stimuli with gain versus loss associations revealed a robust learning asymmetry (Lin, Cabrera-Haro & Reuter-Lorenz, 2020). In the task that Lin et al. (2020) report, participants view a series of neutral image pairs; one pair is associated with a win outcome and another with a loss outcome. Within each pair, one image has a higher probability of leading to a win/loss. Therefore, to maximize rewards participants must learn to select the image that has a high probability of resulting in a win for the win pair and lower probability of resulting in a loss for the loss pair. In other words, participants should learn to make the optimal choice for each pair. Participants showed better learning for the win pair compared to the loss pair, suggesting a learning asymmetry for wins and losses. A meta-analysis of prior reports using this Value Learning Task (VLT) developed by Raymond and O'Brien (2009), or similar probabilistic learning tasks involving gains and losses revealed the prevalence of this learning asymmetry. Furthermore, Lin et al. (2020) demonstrated that this learning asymmetry was robust despite varying whether the win and loss outcomes led to point or monetary earnings, and whether or not participants received detailed instructions about the outcome contingencies.

Nevertheless, within this large data set ($N = 191$), exploratory analyses revealed a subgroup of individuals who, by the final block of trials, had learned the optimal choice for the loss pair better than the win pair (see Supplemental Material). This exploratory analysis raises the question of whether the learning asymmetry evident across the entire sample is universal and a fundamental property of the VLT, or whether there are individual difference variables that influence this pattern. Based on previous research showing individual differences in reward versus punishment learning, it is possible that despite the overall learning asymmetry observed by Lin et al. (2020), there are underlying individual differences in the extent to which gains are learned better than losses.

Previous research has examined individual differences in learning how to earn rewards and avoid punishment by focusing on factors that modulate learning, such as cognitive impulsivity or sensitivity to reward and punishment. However, another approach to examining individual differences is to compare the consistency of learning patterns that people exhibit on different test occasions. That is, to the extent that asymmetries in reward learning vary among individuals and can be attributed to stable individual differences, then, given consistency in the task structure, learning of win and loss associations should be consistent across sessions. Therefore, the present study had two aims: 1) to test the hypothesis that there are underlying individual differences in the extent to which gains are learned better than losses, and 2) to examine whether learning patterns for win and loss associations are consistent within an individual across testing sessions. In order to examine the consistency of learning patterns, we had participants complete the VLT on two separate occasions. To our knowledge, intraindividual consistency in learning patterns in the VLT have not been explored.

Exploring Individual Differences in Value Learning using Model-derived Parameters

One prominent approach that can be used to investigate individual differences in reward learning is based on reinforcement learning theory. Reinforcement learning provides a general framework for studying how artificial and natural systems optimize behavior by learning to predict the consequences of their actions in the environment (Sutton & Barto, 1998). A computational approach is particularly useful for studying individual differences in learning because complex learning behaviors can be captured by a few meaningful parameters. Two informative parameters are the learning rate (α), which describes how much weight is given to new relative to old information, and balance between exploration and exploitation (i.e., inverse temperature parameter, β) (Daw, 2011; Sutton & Barto, 1998). The learning rate ranges from 0 to 1, with an α level of 0 reflecting an agent who learns nothing (exclusively exploiting prior knowledge with no updating), while an α level of 1 reflects an agent who considers only the most recent information (ignoring prior knowledge). In terms of the balance between exploration and exploitation, participants who tend to choose the image with the highest expected value (greater exploitation) will display larger β values, while participants with choice behavior that is less dependent on the image's value (and instead show more exploration) will display smaller β values.

The learning parameters derived from computational reinforcement learning models can be compared between individuals to reveal different learning styles. Previous research that has examined model-derived parameters for learning win and loss associations has demonstrated that behavioral differences in a reinforcement learning type task involving positive and negative outcomes can be attributed to differences in the rate at which participants incorporate feedback from previous trials to guide their choice behavior (e.g., Aberg et al., 2016; van den Bos et al., 2012). For example, using a probabilistic selection task, participants first learned the association between symbols and probabilistic positive and negative outcomes (Aberg et al. 2016). Then, in a subsequent test phase, the symbols from the original trained pairs were mixed up to create novel pairs. Participants used what they learned about the reward probabilities in the training phase to choose the symbol that had a higher likelihood of resulting in a positive outcome and avoid the symbol that had a higher likelihood of resulting a negative outcome. When examining individual differences in performance during the test phase, some participants were better at selecting symbols previously associated with frequent positive outcomes ('approach learners') while others were better at rejecting symbols previously associated with frequent negative outcomes ('avoidance learners'). A computational approach revealed that approach learners showed a slower learning rate (i.e., integrated information across many trials rather than emphasizing the most recent outcomes slowly) for positive outcomes compared to negative outcomes while avoidance learners displayed the reverse pattern. While the probabilistic learning task used in the study by Aberg et al. (2016) differs from the VLT in various ways, such as combining positive and negative outcomes in one pair, and including pairs with different probabilities (80%/20%, 70%/30%, 60%/40%), this example illustrates how a computational approach can elucidate the mechanisms mediating individual differences in learning styles. Therefore, applying reinforcement learning theory to patterns arising in the VLT used by Lin et al. (2020) and similar studies, may provide a computational characterization for individual differences in learning win and loss associations.

In a recent investigation, our team examined the distribution of α and β parameters for participants who learned the win pair better than the loss pair by the end of the task (i.e., win learners) versus those who learned the loss pair better than the win pair better by the end of the task (i.e., loss learners) (Hao et al. 2019). While these parameters are calculated across valence pairs, they can provide insight into whether participants who display different learning patterns in the VLT are characterized by different learning rates and balance between exploration and exploitation. Loss learners had significantly higher α values and lower β values compared to win learners. These results suggest that when making a choice, loss learners updated the expected value of stimuli with feedback from more recent trials and showed greater exploration compared to those who learned wins better than losses. Thus, in the context of the VLT, the extent to which participants incorporate feedback from previous trials and explore the different possible actions can differentiate participants who learn win associations better than loss associations or vice versa. Consistent with previous research, these results demonstrate that parameters derived from reinforcement learning computational models can describe choice behavior of participants displaying different learning patterns in the VLT. However, whether participants' learning rate and choice strategy are stable parameters across multiple instances of learning win and loss associations is unclear.

The present study takes this result a step further by examining the consistency of these parameters across sessions. To do so, the learning rate (α) and balance between exploitation versus exploration (β) parameters were examined in the following ways: 1) whether different learning patterns in the VLT were correlated with the learning rate and inverse temperature parameters, and 2) compared between the two sessions to test the intraindividual consistency of learning rate and choice behavior. Examining whether the reinforcement learning parameters are correlated with performance on the VLT will serve as an extension of the findings by Hao et al. (2019) by demonstrating that behavioral differences in the VLT can be characterized by differences in learning rate and choice strategy. Based on the findings from Hao et al. (2019) we expect that participants who learn losses better than wins will show an aggressive learning rate (higher α) by incorporating feedback from recent trials and a higher degree of exploration (lower β value) compared to participants who learn wins better than losses. Comparing these parameters between the two sessions will reveal the extent to which the learning rate and balance between exploration and exploitation is consistent across time.

Present Study

To summarize, our aim was to test the hypothesis that there are underlying individual differences in the extent to which win associations are learned better than loss associations, and to test the hypothesis that learning patterns for wins and losses are consistent across sessions. To test these hypotheses, participants completed the VLT on two separate occasions. If there are underlying individual differences in learning patterns in the VLT, then learning for win and loss associations should be consistent across the two sessions. On the other hand, it is possible that the learning asymmetry is influenced by a participant's experience with the task that can vary during different task sessions. For example, the learning asymmetry could be a product of the sequence of probabilistic outcomes resulting from participants' choice behavior. When completing the VLT on a second occasion with new stimuli, given that stimulus sequences are randomized, participants will experience a new series of outcomes resulting from their choices. Thus, their learning pattern may differ between the two sessions. By having participants complete the VLT during two sessions, we can assess the extent to which intra-individual consistency characterizes their learning.

As noted, several individual difference variables may be relevant to the learning asymmetry that arises between learned value for wins compared to losses. Previous research has demonstrated that individual differences in trait sensitivity to reward and punishment is associated with the tendency to seek positive outcomes and avoid negative outcomes in behavioral tasks. Therefore, in the present study we also investigated whether the psychological constructs of motivation, impulsivity, and sensitivity to punishment and reward were associated with learning patterns for win and loss associations in the VLT. The results of this research can provide insight into psychological traits that may correlate with learning of win and loss associations. In addition, we applied reinforcement learning theory to elucidate the computational mechanisms mediating individual differences in learning patterns and consistency in learning rate and choice strategy. The results will build on the findings from Hao et al. (2019) in characterizing different learning patterns based on participants' choice behavior on a trial-level basis.

To gain further insight into what people learn about value, we included a memory task that probed participants' explicit memory of the outcomes associated with the images they viewed in the VLT. The results of the memory task could provide clues about the memory representations that are formed and whether there are individual differences in the explicit knowledge people have about win- and loss-associated stimuli that can be linked to variability in learning patterns in the VLT.

Method

Participants

Sixty-nine young adults were recruited from the University of Michigan for a two-part study and received $10 per hour for their participation. The minimum number of participants required was determined by an *a priori* power analysis (G*power; Faul, Erdfelder, Buchner, & Lang, 2009) with additional participants recruited to account for participants who might fail to reach the learning criterion and to allow us to explore individual differences. Data were excluded from six participants who did not return for the second session and three additional participants due to computer malfunction that occurred during the VLT. The remaining 60 participants (49 females) had a mean age of 19.57 years (range, 18 – 25 years). All participants gave written informed consent for the study, which was approved by the University of Michigan, Institutional Review Board. All participants were right-handed and were prescreened to exclude those who had a history of depression, anxiety, ADHD, head injury, or were currently taking medications that affect cognition. Following the procedure used in several prior studies of the VLT, we adopted a learning criterion for the win and loss conditions (O'Brien & Raymond, 2012; Painter, Kritikos, & Raymond, 2014; Rutherford, O'Brien, & Raymond, 2010). We excluded participants who did not select the optimal choice on at least 65% of the trials for both win and loss conditions in the final block of the VLT during both sessions (n = 4)[2]. Therefore, the final

---

[2] Note that the cutoff point for achieving learning is at the low end of performance. The pattern of results remained the same when including all participants.

sample size for data analysis was 56 participants, (46 females; M_{age} of 19.57, with age ranging 18-25 years).

Materials and Procedure

Participants completed tasks over the course of two sessions, with the second session occurring one week after the first session (Figure 2.1). During the first session, participants completed a series of questionnaires and the VLT. Completion of questionnaires lasted approximately 25 minutes, while completion of the VLT lasted approximately 20-30 minutes. During the second session, participants completed a second version of the VLT, with new stimuli that were not presented in the first session, and a memory task consisting of images that were presented in the VLT during the second session and new images that did not appear in either the first or second session. Completion of the VLT lasted approximately 20-30 minutes, while completion of the Memory Task lasted approximately 5 minutes.

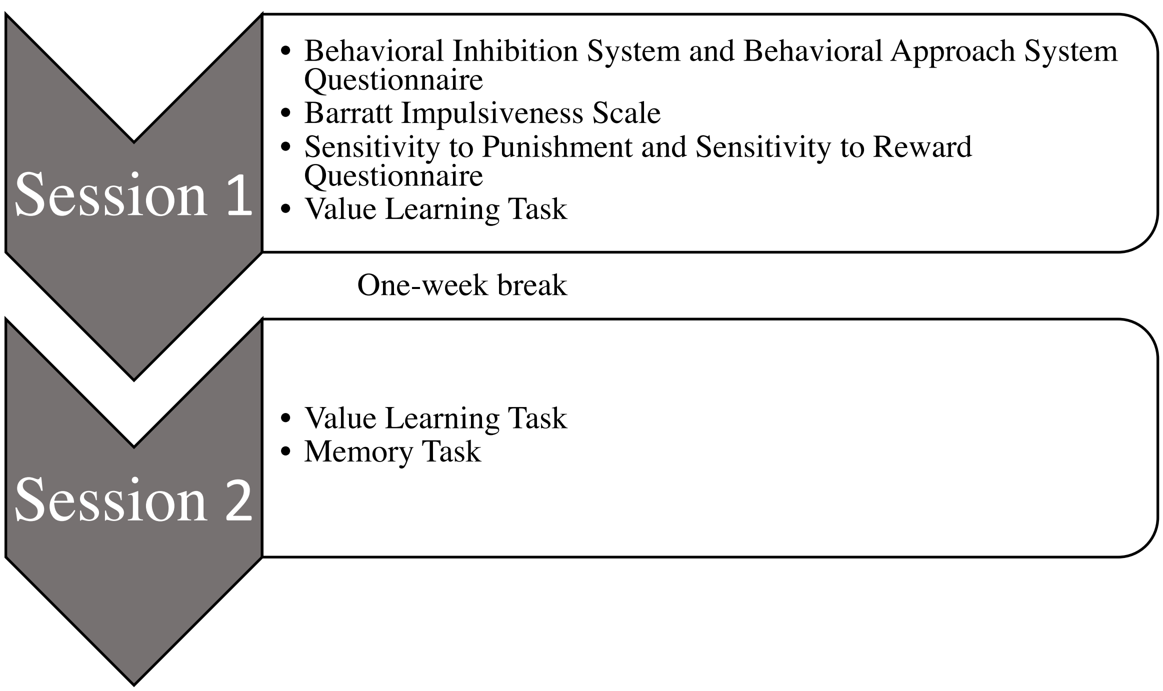


Figure 2.1. Study protocol for each session.

Session 1 lasted one hour. Session 2 lasted 30 minutes.

Questionnaires. A measure of the two general motivational systems underlying behavior was obtained using the Behavioral Inhibition System and Behavioral Approach System (BIS/BAS) questionnaire (Carver & White, 1994). This 24-item questionnaire asks participants to rate the extent to which they agree or disagree with each statement on a 4-point scale (1 = *very true of me*, 4 = *very false of me*). Scores are created by summing the items corresponding to each subscale (BAS Drive, BAS Fun Seeking, BAS Reward Responsiveness, and BIS).

A measure of the personality and behavioral construct of impulsivity was obtained using the Barratt Impulsiveness Scale (BIS-11; Patton, Stanford, & Barratt, 1995). This 30-item questionnaire consists of statements describing common impulsive or non-impulsive behaviors and preferences. Participants rate how often they engage in each behavior on a 4-point scale (1 = *rarely/never*, 4 = *almost always/always*). Scores are created by summing the items corresponding to each subscale (Attentional Impulsivity, Motor Impulsivity, and Non-Planning Impulsivity).

A measure of how individuals respond to reward compared to punishment was assessed using the Sensitivity to Punishment and Sensitivity to Reward questionnaire (SPSRQ; Torrubia, Avila, Molto, & Caseras, 2001). This 48-item questionnaire consists of statements describing behaviors that an individual may engage in. Participants indicated whether the behaviors are descriptive of them by circling either "*yes*" or "*no.*" Scores are created by adding all the "yes" answers to items corresponding to the Sensitivity to Punishment and Sensitivity to Reward subscales.

Stimuli. Twenty-four grayscale images of landscape scenes were used as the visual stimuli (6 for the VLT during the first session, 6 for the VLT during the second session, 12 for the memory task). The images were acquired from the Cognitive Neuroscience Research Lab led by Adam Gazzaley (Rissman, Gazzaley, & D'Esposito, 2009). The assignment of scenes to the VLT during the first and second session was counterbalanced across participants. Images were 225 pixels wide x 225 pixels tall. All stimuli were presented using E-prime 2.0 (Schneider, Eschman, & Zuccolotto, 2002) on a 40-cm monitor (75-Hz refresh, 1152 x 864 resolution) of a Pentium 4 computer. Viewing distance was approximately 45 cm.

Value Learning Task. After the questionnaires, participants completed the VLT, which was a modified version of the task used by Raymond and O'Brien (2009). Six scenes were divided into three pairs each of which was assigned to one of three valence conditions: win, loss, or no change. The pairing of scenes, assignment of each scene to a valence and salience condition, and assignment of each scene to either session 1 or session 2 was counterbalanced across participants. The task consisted of 5 blocks of trials, with each of the 3 pairs (win pair, loss pair, and no change pair) presented 20 times per block in a random order, yielding a total of 300 trials across all pairs. Participants were given a one-minute break between blocks. On each trial, a scene pair was presented on a computer screen, one above and one below a central fixation cross (Figure 2.2). Each scene was always presented with its pair, but the location of the scene (top vs. bottom) was randomized from trial to trial. For each trial, participants selected one scene from the pair, with the goal to maximize the number of points they earned. Participants pressed the "f" key to select the top scene or "j" key to select the bottom scene. For the win pair, selection of a scene resulted in either a win or no change; for the loss pair, selection of a scene resulted in either a loss or no change; for the third pair (no change pair), selection of either scene resulted in no change in points. The no change pair served as a control to examine whether there were any systematic biases in the tendency for participants to select the top or bottom image.

Figure 2.2. Schematic of the Value Learning Task.
Participants view a pair of neutral landscape scenes, one above and one below a central fixation cross. Participants select one scene from a pair by pressing one of the two designated keys on the computer keyboard. A yellow border appeared on the scene selected, remaining for 500 msec. Then, the outcome of participants selection and the resulting score was displayed on the screen for 1500 msec. The ITI was 500 msec. In this example, a participant selected a scene that resulted in a win.

The probability of the outcome (high vs. low) also varied for the scenes constituting the win and loss pairs. For each win pair and loss pair, one scene produced the corresponding win or loss outcome with a probability of .80 (high likelihood) while its counterpart produced its assigned win or loss outcome with a probability of .20 (low likelihood). The assigned valence and the contingencies were constant for the entire 5-block session. To maximize points, participants had to learn which scene led to a higher probability of gaining points (80% win scene) and which scene led to a lower likelihood of losing points (20% loss scene). Thus, 'optimal selection' was defined as selecting the 80% win scene when the win pair was displayed and 20% loss scene when the loss pair was displayed.

Immediately after the participant chose a scene, the computer screen displayed the message “Win” in green, “Lose” in red, or “No change” in black, depending on the scene pair just presented and the probability governing the outcome. A win or a loss outcome resulted in a gain or loss of 5 points, whereas no change resulted in no change in points. A running total of the earnings appeared in the top right-hand corner following each selection. Prior to beginning the task, participants were told that they should earn as many points as they could.

Memory Task. During the second session, once participants completed the VLT they then performed the Memory Task. This task was designed to test participant’s explicit knowledge of the outcome associated with each scene from the VLT. The Memory Task consisted of the 6 scenes shown in the VLT from the second session along with 12 new scenes that did not appear in either the first or second session. Each scene was presented one at a time at the center of the computer screen. The scenes from the VLT were each presented four times, whereas new scenes were each presented two times in a randomized order yielding a 50% probability of the old versus new items. For each scene, participants indicated which outcome the scene was most likely to be associated (very likely to lead to a win, occasional win, no change, occasional loss, or very likely to lead to a loss), or whether it was a new image (Figure 2.3). Completion of the memory task was self-paced, and no feedback was given about performance on this task.

1 = Very likely to win, 2= Occasionally win,
3 = No change,
4 = Occasionally lose, 5 = Very likely to lose,
6 = None (new image)

Figure 2.3. Schematic of the Memory Task.
This task consists of the 6 scenes from the Value Learning Task and 12 new scenes. Participants select the outcome that was most likely associated with each scene.

Results

Value Learning Task

For the purpose of comparing learning for scenes associated with wins or losses, the probability of optimal choice was only examined for the win and loss pairs. Performance on the VLT was examined by calculating the probability of optimal choice for each pair: the proportion of trials for which participants chose the high-probability win scene (80% win) for the win pair and low-probability loss scene (20% loss) for the loss pair. The probability of optimal choice was examined for bins of 20 trials, resulting in five 20-trial blocks for each pair. Mean probability of optimal choice for the win pair and loss pair were then compared using a 5 (Trial Block: 1-5) x 2 (Outcome Valence: win, loss) x 2 (Session: Session 1, 2) within-subjects ANOVA. Greenhouse-Geisser corrections were applied when assumptions of sphericity were violated, and *p*-value significance cutoffs were adjusted using a Bonferroni correction (α /number of comparisons) to account for multiple pairwise comparisons.

The results revealed a significant main effect of Trial Block indicating that participants' probability of selecting the optimal scene differed as they progressed through the task, $F(2.42, 137.14) = 75.08$, $p < .001$, $n_p^2 = .58$. Follow-up pairwise comparisons revealed a significant increase in the probability of optimal choice across Trial Blocks, *p*s < .01, except between Block 3 and Block 4, $p = .34$. A significant main effect of Outcome Valence emerged, indicating that participants' probability of selecting the optimal scene was higher for win scenes ($M = .81$, $SE = .02$) compared with loss scenes ($M = .78$, $SE = .02$), $F(1, 55) = 4.67$, $p = .04$, $n_p^2 = .08$ (Figure 2.4). No significant main effect of Session emerged indicating that overall, learning did not significantly differ between the two sessions, $F(1, 55) = 2.45$, $p = .12$. A significant interaction between Trial Block and Session emerged, $F(2.66, 146.00) = 3.05$, $p = .04$, $n_p^2 = .05$. Follow-up pairwise comparisons revealed that during Session 1, participants' probability of selecting the optimal scene increased between Blocks 1 through Block 3, *p*s < .001, but did not increase between Blocks 3 through Block 5, *p*s > .46. During Session 2, participants' probability of selecting the optimal scene increased between Block 1 and 2, $p < .001$, but did not differ between any of the subsequent blocks, *p*s > .05. Thus, participants' learning of the optimal outcome reached an asymptote earlier during Session 2 when compared with Session 1. No significant interactions emerged between Outcome Valence and Trial Block, Outcome Valence and Session, or between Outcome Valence, Trial Block, and Session, *p*s > .13.

For the no-change pair, neither scene is an optimal choice. Therefore, one no-change scene was arbitrarily selected as the correct scene. The probability of selecting the scene was examined by a repeated-measures ANOVA as a function of Trial Block (1-5) and Session (Session 1 vs. 2). The probability of selecting the chosen "correct" no-change scene was .48 across all participants and did not significantly differ by Trial Block, Session, or Trial Block x Session, Fs < 2.10, ps $> .10$. Thus, there were no systematic biases in selecting the arbitrarily selected correct or incorrect scene.

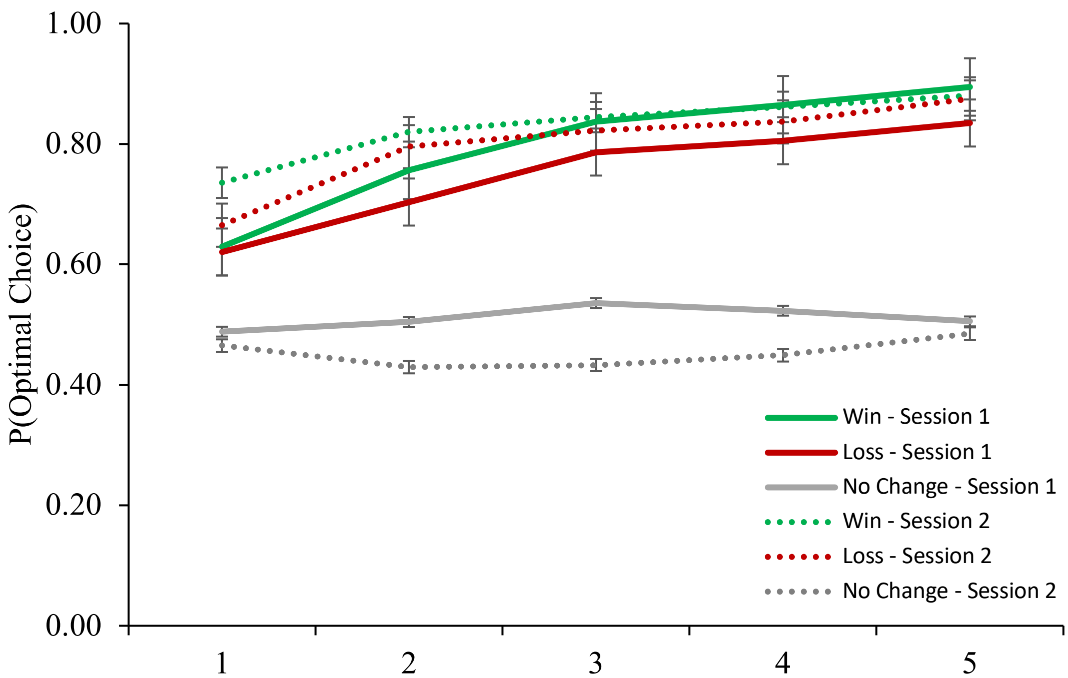


Figure 2.4. Value Learning Task Performance.
The probability of participants selecting the optimal scene for the win pair was significantly higher than for the loss pair. This pattern did not differ significantly between Session 1 and Session 2. Participants' learning of the optimal outcome reached an asymptote earlier during Session 2 when compared with Session 1.

Win/Loss Learning Difference Across Sessions. To examine whether the learning difference between win and loss associations was consistent across time, a learning difference score was calculated for each individual during each session by subtracting the percent of optimal choice for the loss pair from the percent of optimal choice for the win pair during the final block of trials (i.e., last 20 trials). The shift in difference scores for each participant is demonstrated in Figure 2.6. A Pearson correlation was conducted between the two difference scores. The results revealed that the difference score for Session 1 ($M_{difference}$ = 0.06, *SD* = .17) was not correlated with the difference score for Session 2 ($M_{difference}$ = .01, *SD* = .20), $r(56) = -.10$, $p = .47$ (Figure 2.5).[3] Thus, this measure provided no evidence for consistency in the learning asymmetry between the two sessions.[4] It is possible that the likelihood of a significant correlation between the learning difference scores was reduced due to no correlation between win learning during Session 1 and 2 and/or loss learning during Session 1 and 2. Therefore, we explored whether probability of optimal choice was correlated across sessions separately for the win pair and loss pair. The results showed that probability of optimal choice for the win pair was not correlated between Session 1 and Session 2, $r(56) = .18$, $p = .18$ (Figure 2.7) but probability of optimal choice for the loss pair for Session 1 and Session 2 were significantly correlated, $r(56) = .35$, $p = .01$ (Figure 2.8).[5] In other words, while participants who had higher optimal choice behavior for the loss pair during Session 1 also tended to have higher optimal choice behavior for the loss pair during Session 2, win learning was not correlated across sessions. Therefore, the non-significant correlation for win learning between Session 1 and 2 may have contributed to the non-significant correlation between the learning difference scores. It is possible that the lack of a significant correlation for win learning was due to lower variability, given that win learning

[3] Note that when examining the learning difference between win and loss pairs across all five blocks of trials, the correlation remained non-significant, $r(56) = -0.05$, $p = .70$.
[4] A Bayes factor test for the correlation indicated that the data are 2.62 times more likely under the null hypothesis (that there is non-significant correlation) compared to the alternative hypothesis.
[5] When examining probability of optimal choice across all five blocks the pattern remained the same with no significant correlation for win accuracy across sessions, $r(56) = 0.03$, $p = .86$, and a significant correlation for loss accuracy across sessions, $r(56) = .55$, $p < .001$.

tends to be at ceiling. To test this possibility, we computed Mauchly's Test of Sphericity to compare the variances between win learning and loss learning for each session. The results revealed that equal variances could not be assumed ($p < .001$).

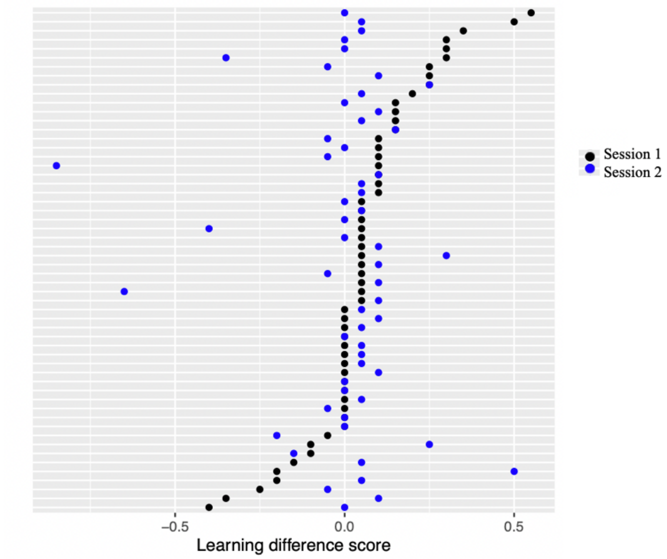


Figure 2.6. *Learning difference scores for each participant, sorted by the learning difference score during Session 1.*

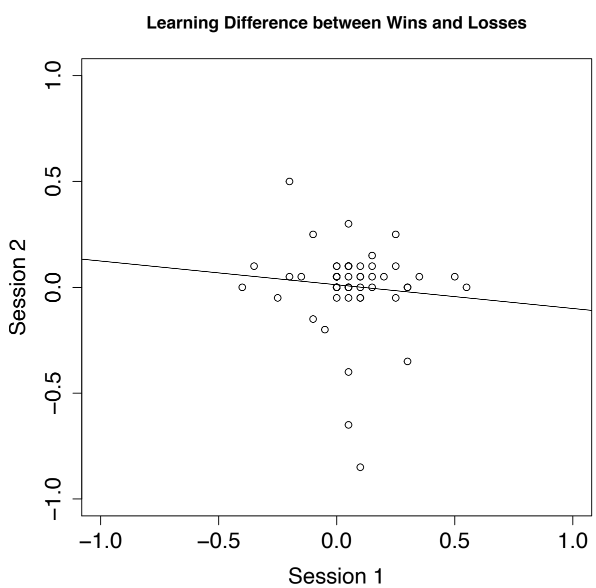


Figure 2.5. *Scatterplot for the Learning Difference Score between Session 1 and 2.*

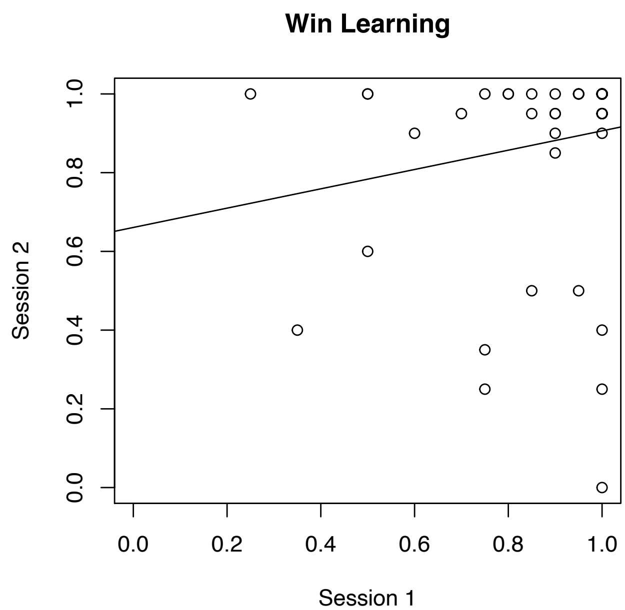


Figure 2.7. Scatterplot for win learning between Session 1 and Session 2.

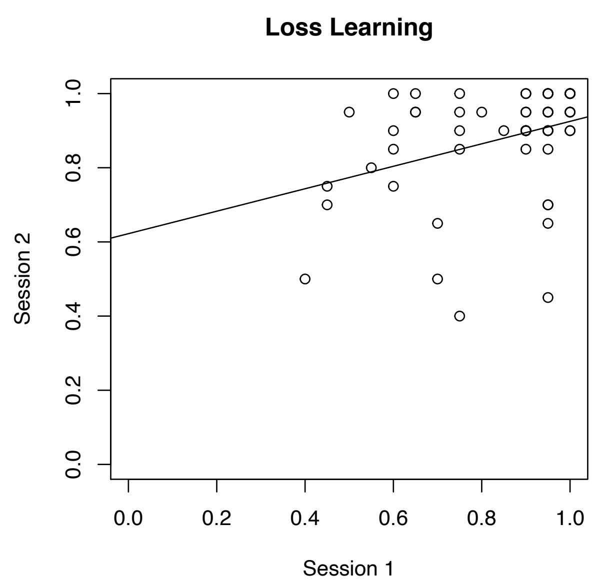


Figure 2.8. Scatterplot of loss learning between Session 1 and 2.

Correlations between Questionnaire Measures and Performance on the VLT and RL Parameters. To examine whether measures of behavioral inhibition and avoidance, impulsivity, and sensitivity to punishment and reward are associated with learning win and loss outcome associations (average percent of optimal choice during the final block of trials), a series of correlations were conducted. A Holm-Bonferroni correction method was used to adjust the alpha level for multiple comparisons. The results showed a significant correlation for loss learning during Session 2 and the measure of attentional impulsivity such that participants with greater impulsivity tended to have lower levels of loss learning (Table 2.1).

We also examined whether the questionnaire measures were associated with the reinforcement learning parameters. A Holm-Bonferroni correction method was also used to adjust the alpha level for multiple comparisons. The results showed no significant correlations between win and loss learning during Session 1 and 2 and the learning rate and inverse temperature parameters (Table 2.2).

Table 2.1.
Pearson r Values for the Correlations between Questionnaire Measures and VLT Performance.

Percent Optimal Choice	BAS Drive	BAS Fun	BAS Reward	Behavioral Inhibition	Attentional Impulsivity	Motor Impulsivity	Non-planning Impulsivity	Sensitivity to Punishment	Sensitivity to Reward
Win Pair (Time 1)	.00	.09	.03	.20	-.13	-.22	-.15	.12	-.07
Lose Pair (Time 1)	-.05	.14	.06	.04	-.22	-.21	-.23	.07	-.09
Win Pair (Time 2)	.20	.06	.09	.01	-.23	-.13	-.07	-.08	-.03
Lose Pair (Time 2)	.06	.10	-.03	.09	-.37**	-.20	-.17	-.12	-.10

Note. BAS = Behavioral Approach System.
**p* < .05, ***p* < .01, ****p* < .001

Table 2.2.
Pearson r Values for the Correlations between Questionnaire Measures and RL Parameters.

Percent Optimal Choice	BAS Drive	BAS Fun	BAS Reward	Behavioral Inhibition	Attentional Impulsivity	Motor Impulsivity	Non-planning Impulsivity	Sensitivity to Punishment	Sensitivity to Reward
α -Time 1	-.08	-.05	-.10	.02	.06	.03	.11	.05	.08
β - Time 1	.02	.14	.10	-.11	-.30	-.17	-.32	.11	-.03
α - Time2	.08	-.16	-.10	.12	-.27	-.03	-.07	-.25	-.11
β - Time 2	-.13	.23	.01	.21	-.09	-.16	-.14	-.00	.05

Note. α is the learning rate, β is the inverse temperature parameter.
$*p < .05$, $**p < .01$, $***p < .001$

Memory Task

To examine participants' explicit knowledge of the outcomes associated with each VLT scene, a binary logistic regression analysis was performed to examine the odds of correctly indicating the outcome associated with each VLT scene (accurate outcome coded as 1, incorrect outcome coded as 0). The predictor variables were valence (win versus loss, with win coded as 0), optimality (suboptimal versus optimal, with suboptimal coded as 0), and interaction between valence and optimality. A test of the binary logistic regression model with the interaction term versus without the interaction term was not statistically significant, indicating that including the interaction term between valence and optimality did not increase model fit, χ^2 (1, N = 56) = 0.00, p = 1.00. Therefore, the final model only included valence and optimality as potential predictors. Results of the binary logistic regression revealed that optimality was the only significant predictor in the model indicating that the odds of participants correctly selecting the outcome for optimal scenes (80% win for the win pair and 20% loss for the loss pair) were higher than for suboptimal scenes (20% win for the win pair and 80% loss for the loss pair) ($p < .001$; Figure 2.9).

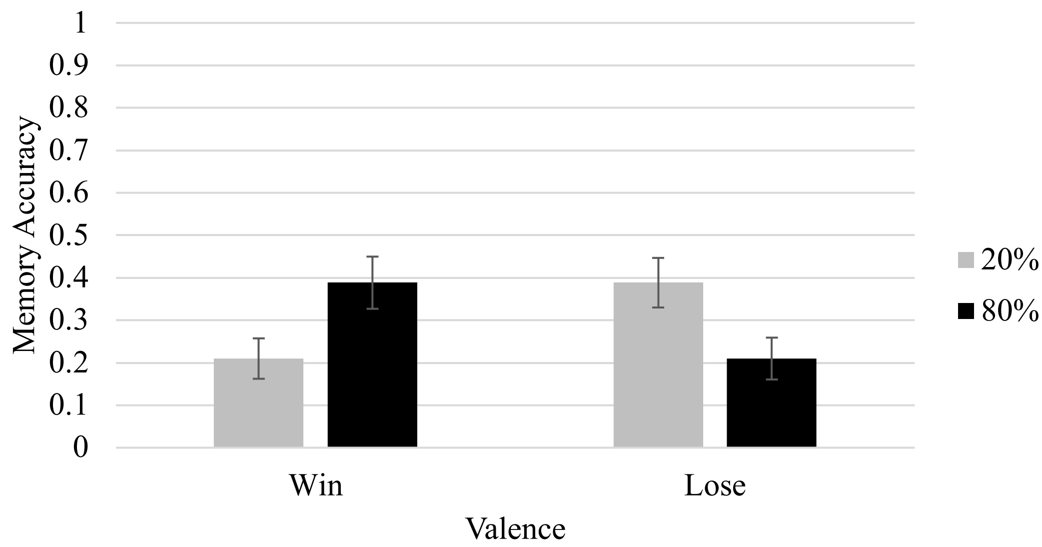


Figure 2.9. Memory Task Performance.
The odds of participants correctly selecting the outcome of optimal scenes (80% win scene for the win pair; 20% loss scene for the loss pair) were higher than for suboptimal scenes (20% win scene for the win pair; 80% loss scene for the loss pair).

We also examined the full range of participants' responses by conducting a multinomial logistic regression with valence, probability, and the interaction between valence and probability as predictors, and all response options for the memory task as the outcome variable (Table 2.3). The response option "No Change" was selected as the reference level.

Table 2.3.
Multinomial Logistic Regression Predicting Outcome Response for each VLT scene from Valence and Probability (N = 56)

Outcome Response	Valence	Probability	Valence * Probability
Very likely to win	.00	0.95	4.43×10^5
Occasionally win	.17***	0.90	2.59
Occasionally lose	7.01***	0.76	1.54
Very likely to lose	23.70***	4.82*	1.58*
None (new)	1.38	0.59	6.46

Note. Values in the table are odds ratio. Valence was coded as 0 for the win scenes, and 1 for the loss scenes. Probability was coded as 0 for low probability, and 1 for high probability.
p* < .05, ** *p* < .01, * *p* < .001

For win scenes, probability was not a predictor of outcome attribution. In other words, outcome attribution responses did not significantly differ between the 20% win scene and 80% win scene. “Very likely to win” was most frequently selected as the outcome associated with the scenes, followed by “no change” and “occasionally win” (Figure 2.10). Valence misattribution was rare. Thus, for win scenes, attribution errors were primarily confusing the probability that a win outcome occurred.

For loss scenes, outcome attribution responses did not significantly differ between the 20% loss scene and 80% loss scene (Figure 2.11). “Occasionally lose” was most frequently selected as the outcome associated with the scenes, followed by “very likely to lose” and “no change.” Participants rarely attributed the incorrect valence to loss scenes. Thus, similar to win scenes, attribution errors for loss scenes were confusing the probability that a loss outcome occurred.

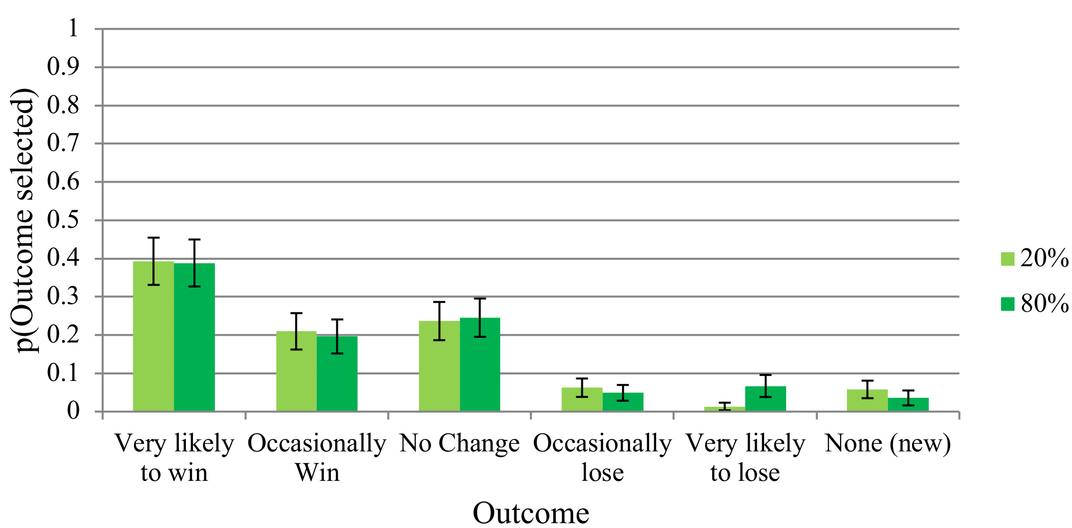


Figure 2.10. Outcome Attributions for Win Scenes.

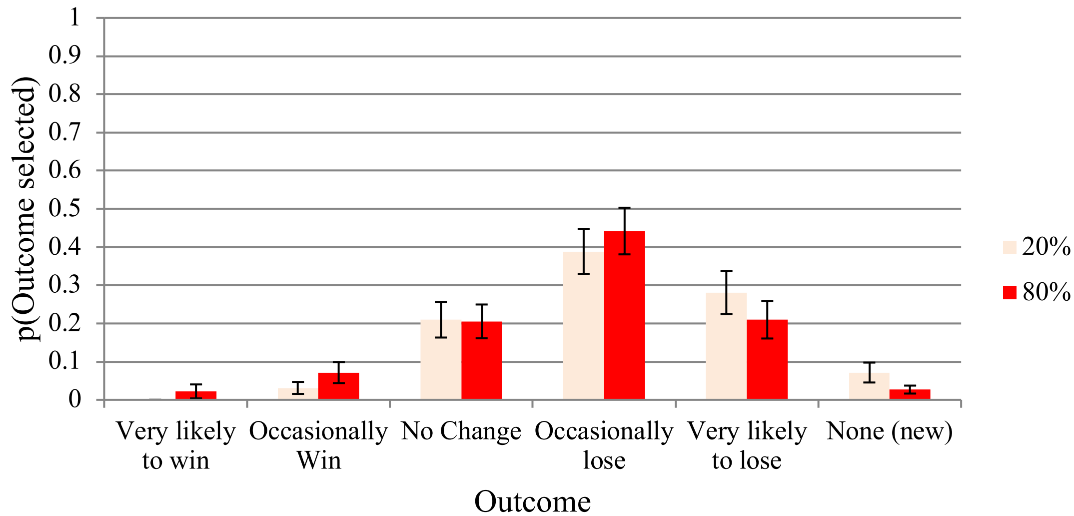


Figure 2.11. Outcome Attributions for Loss Scenes.

Correlation Between Value Learning and Memory Accuracy for Win and Loss Scenes

We also examined the relationship between value learning performance for win and loss scenes in session 2 and accuracy for the corresponding win and loss scenes. A correlation for each image (80% win, 20% win, 80% loss, 20% loss) was computed and the significance was determined after correcting for multiple correlations using a Holm-Bonferroni correction. With the correction, no significant correlations emerged between win learning in the VLT and memory accuracy for either win scene (20%: $r(54) = -.24$, $p = .07$; /80%: $r(54) = .25$, $p = .07$). No significant correlations emerged between loss learning in the VLT and memory accuracy for either loss scene (20%: $r(54) = .02$, $p = .91$/80%: $r(54) = .09$, $p = .51$).

Reinforcement Learning

We used reinforcement learning model to derive the learning rate (α) and choice strategy (β) parameters for each participant based on their choice behavior in the VLT. A Softmax rule and maximum likelihood estimation was implemented to find the pair of parameters that would most likely coincide with the same choice on each trial as each participant (Daw, 2011). The Q-learning update rule was used to estimate α such that:

$$Q_{t+1}(a) = Q_t(a) + \alpha(r_t - Q_t(a)),$$

where $Q(a)$ represents the expected reward when action α is selected, $Q_t(a)$ is the estimated value of action a at trial t, and α is the agent's learning rate with values ranging between 0 and 1.

The Softmax choice rule was used to convert Q-value estimates into choices while balancing exploration and exploitation. According to the Softmax rule, at trial t, the probability of choosing an action A given the value estimates for action A and B is:

$$P\big(c_t = A \big| Q_t(A), Q_t(B)\big) = \frac{\exp\,(\beta * Q_t(A))}{\exp\big(\beta * Q_t(A)\big) + \exp\big(\beta * Q_t(B)\big)}$$

where β is the inverse temperature parameter accounting for the extent to which participants explore different options versus exploit what they have learned. The search for the best-fitting parameters was done with each combination of $\alpha\ \epsilon[0,1]$ with step size 0.01 and $\beta\ \epsilon$ [0,10] with step size 0.1. For each participant, we chose the pair of parameters that produced the largest log-likelihood as the maximum likelihood estimation of that individual's learning rate and choice strategy (Session 1: Figure 2.12; Session 2: Figure 2.13)

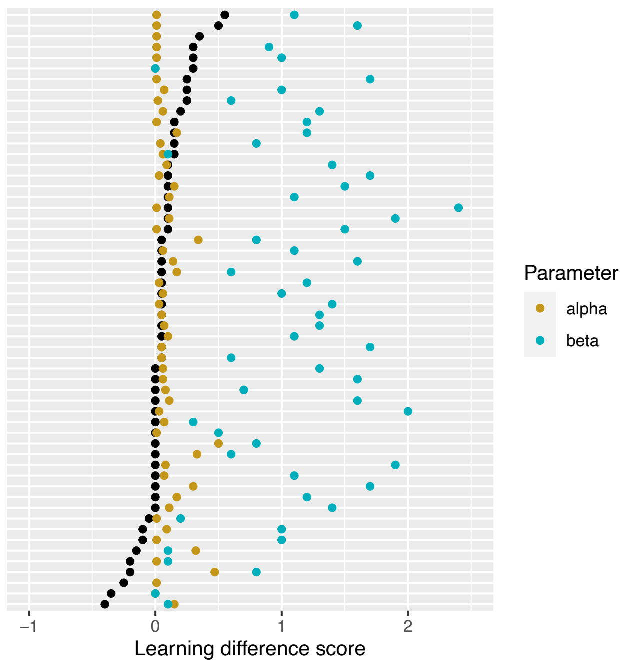


Figure 2.12. Learning difference score and reinforcement learning parameters for each participant during Session 1.

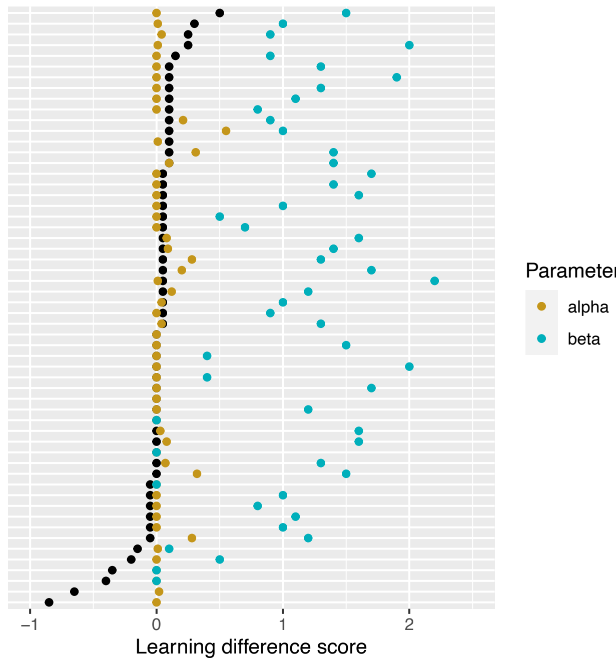


Figure 2.13. Learning difference score and reinforcement learning parameters for each participant during Session 2.

The first goal of examining these reinforcement learning parameters was to test whether participants displaying different learning patterns in the VLT could be characterized by differences in learning rate and choice strategy. To test this, a series of correlations were conducted between each reinforcement learning parameter (α and β) and the learning difference during the final block of trials between wins and losses for each session (Figure 2.14). For Session 1, a more aggressive learning rate (i.e., higher α; $M = .09$, $SD = .11$) was associated with a smaller learning difference between wins and losses, $r(56) = -.30$, $p = .03$. Whereas, the balance between exploration and exploitation (β; $M = 1.17$, $SD = .77$) was not associated with the learning difference score, $r(56) = .14$, $p = .29$. For Session 2, the learning rate ($M = .15$, $SD = .15$) was not associated with the learning difference score, $r(56) = .16$, $p = .24$. Whereas, greater exploitation (higher β; *M* = 1.40, *SD* = 1.18) was associated with a smaller learning difference between wins and losses, $r(55) = -.37$, $p = .005$.

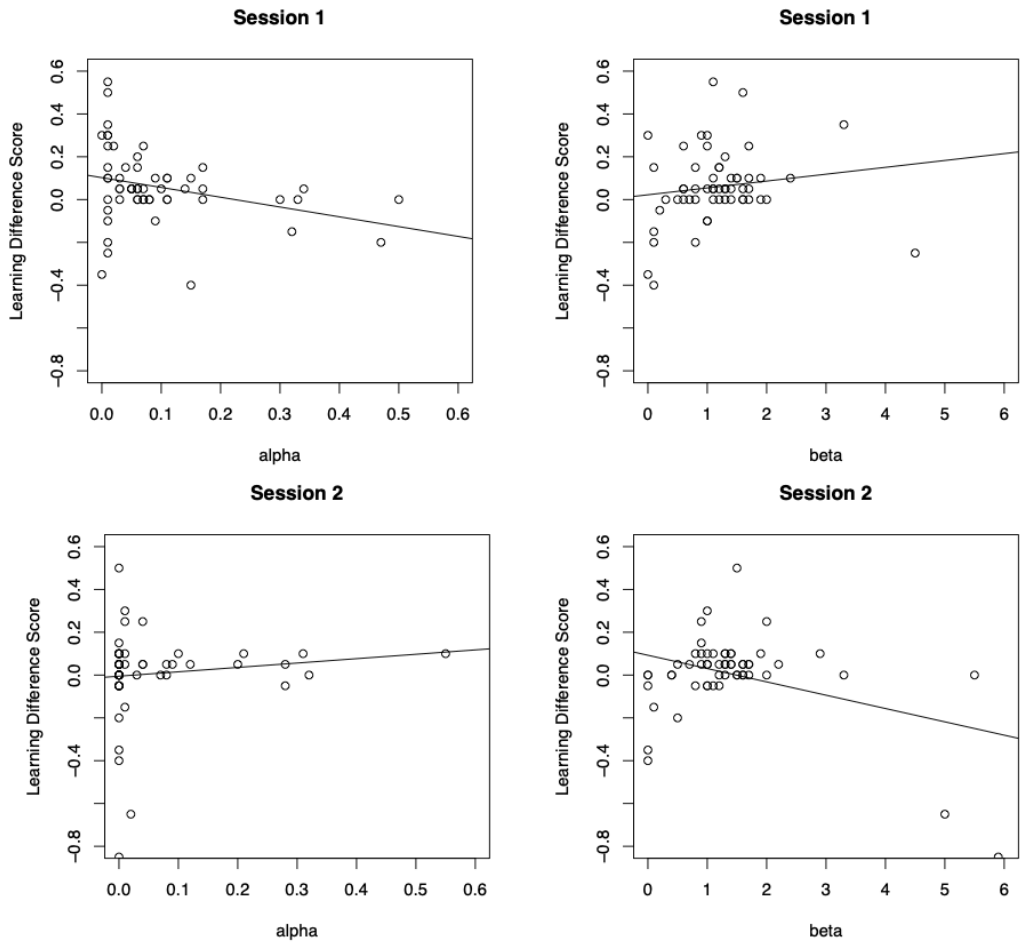


Figure 2.14. Scatterplots of model-derived parameters and the learning difference score. The learning difference score was calculated as the optimal choice for the win pair – optimal choice for the loss pair for the final block of trials in the VLT.

The second goal was to examine the consistency of the reinforcement learning parameters. A correlation was conducted between the learning rate for Session 1 and 2 and the inverse temperature parameter for Session 1 and 2. The results showed that the learning rate during Session 1 and Session 2 were correlated, $r(56) = .28$, $p = .04$ (Figure 2.16). Participants displaying rapid updating of expected value during the first instance of learning tended to also show an aggressive learning rate during the second session. However, the inverse temperature parameter (β) at Session 1 and Session 2 did not correlate, $r(56) = .13$, $p = .35$ (Figure 2.15). Thus, there only appeared to be intra-individual consistency in the learning rate.

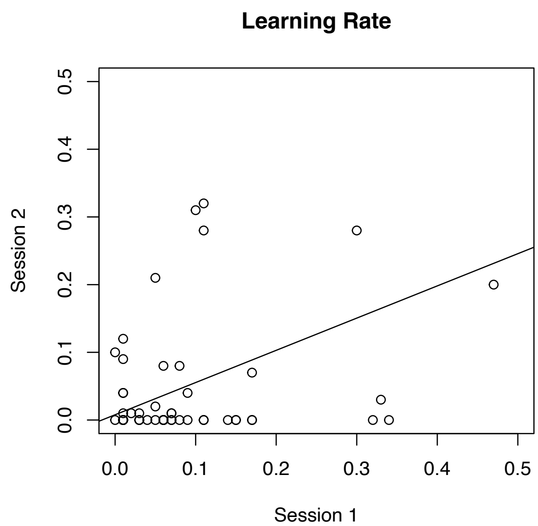


Figure 2.16. Scatterplot of the learning rate across sessions.

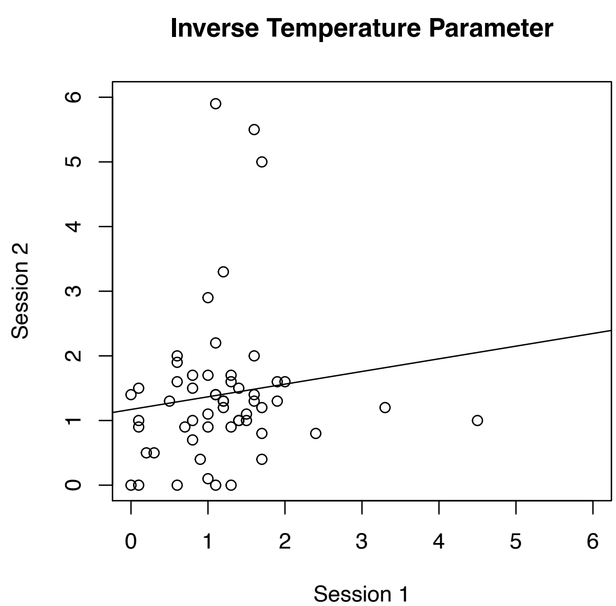


Figure 2.15. Scatterplots of the inverse temperature parameter across sessions.

Discussion

A recent investigation of how people learn to associate neutral stimuli with win and loss associations revealed a learning asymmetry such that participants learned win associations better than loss associations (Lin et al. 2020). However, exploratory analyses of the data set from Lin et al. (2020) revealed a subgroup of individuals who learned losses better than wins in the final block of the task. This finding raised the question of whether the learning asymmetry favoring win associations is a universal pattern or whether individuals differ in the extent to which they learn gains better than losses. Therefore, the goal of the present study was to examine the possibility of underlying individual differences in the extent to which win associations are learned better than loss associations, and whether learning patterns for win and loss associations are consistent within-individuals across sessions. To this end, participants completed the Value Learning Task (VLT), a probabilistic task involving wins and losses, on two separate occasions. During the first session, participants completed the VLT and a series of questionnaires designed to measure behavioral inhibition and approach behaviors, impulsivity, and sensitivity to reward and punishment. During the second session, participants completed the VLT with new stimuli, and a memory task designed to measure participants' explicit knowledge of the outcomes associated with the VLT images.

Value Learning Task

The results from the VLT demonstrated that the learning asymmetry was prevalent during both sessions. The learning pattern, averaged across all participants, was consistent with the previously reported asymmetry showing that participants learned the optimal choice for the win pair better than the loss pair (Lin et al. 2020). When comparing learning between the two sessions, participants showed faster learning during Session 2. This latter finding suggests that participants may acquire general knowledge of the task structure or adopt a different strategy that leads them to learn the optimal choice more quickly during the second instance of learning. However, additional experience with the task was not sufficient to eliminate the learning asymmetry.

Previous research has found that variability in the learning of win and loss associations is modulated by individual differences in variables such as trait sensitivity to reward and punishment (e.g., Kim et al. 2015), cognitive impulsivity (e.g., Cáceres & San Martín, 2018), and approach and avoidance tendencies (e.g., Aberg et al. 2016). In the present study, we took several approaches to examining whether there are individual differences in the extent to which win associations are learned better than loss associations. First, we examined whether the learning difference between wins and losses for Session 1 and 2 were correlated. However, no significant correlation emerged. When examining the correlation between win accuracy across sessions separately from loss accuracy, we found that loss accuracy for Session 1 and 2 were correlated while win accuracy was not correlated between Session 1 and 2. Thus, the non-significant correlation between the learning difference scores may be due to the non-significant correlation in performance for the win pair. It is possible that the lack of a significant correlation for win learning was due to lower variability, given that win learning tends to be at ceiling. To test this possibility, we computed Mauchly's Test of Sphericity to compare the variances between win learning and loss learning for each session. The results revealed that equal variances could not be assumed ($p < .001$). We also examined whether measures of behavioral approach and avoidance behaviors, impulsivity, and sensitivity to reward and punishment were associated with learning for wins and losses but found no significant correlations. Thus, when evaluating behavioral performance of the VLT, there was no evidence that variability in the learning pattern for win and loss associations could be attributed to individual difference variables or that there is intraindividual consistency in learning patterns.

Reinforcement Learning Model Parameters

Reinforcement learning theory provides another approach to examining learning from gains and losses by characterizing choice behavior with meaningful parameters. We derived the learning rate and inverse temperature parameters (i.e., balance between exploration and exploitation) for each participant at both sessions. We first examined whether the degree to which participants incorporated feedback from each trial and their balance between exploration and exploitation were correlated with the learning difference between win and loss associations. We found that during the first session, a more aggressive learning rate was associated with a smaller learning difference between wins and losses. However, during the second session, the learning rate was not correlated with the learning difference score. One potential explanation for this latter finding is that by the second session, there is less variability in the learning difference between wins and losses. While the analysis of VLT performance showed that the learning asymmetry persisted across both sessions, quantitatively, the learning asymmetry was less robust during the second session. Lower variability masks potential correlations between the learning rate and learning difference score. The inverse temperature parameter was not associated with the learning difference score for Session 1, but greater exploitation was associated with a smaller learning difference between wins and losses at Session 2. Thus, participants who showed more similar learning between wins and losses tended to exploit expected value more so than participants who had a larger learning difference between wins and losses. One potential explanation for the significant correlation at Session 2 but non-significant correlation at Session 1 is that those who showed more similar learning during the second session were more likely to base their choice behavior on the knowledge they acquired about the task structure from the first session.

Next, we examined the consistency of the reinforcement learning parameters. The results showed that the learning rate during Session 1 and 2 were correlated while the inverse temperature parameters between Session 1 and 2 were not correlated. These findings suggest that participants who had a more aggressive learning update rule during the first session tended to also update the expected value from recent trials during the second session. In terms of the inverse temperature parameter, participants' balance between exploration versus exploitation was not consistent across the two sessions. Thus, while the VLT showed no strong evidence of intraindividual consistency, the reinforcement learning parameters show evidence of some consistency in choice strategy.

The present study took a different methodological approach to examining learning patterns in the VLT compared to Hao et al. (2019). In Hao et al. (2019), learning patterns were explored by creating categories of learners, whereas in the present study, we used a continuous measure of the learning difference between wins and losses. The limitations of categorizing continuous variables are well known and include losing precision of estimated means or relying on arbitrary cutoff points to create the categories that may not generalize to other data sets (e.g., Altman, & Royston, 2006). Thus, analyses from future data sets can be compared with the current results without the limitation of having arbitrary cutoff points that may be specific to a certain data set.

Explicit Knowledge of Learned Win and Loss Associations

We also assessed participants' explicit knowledge of the outcomes associated with each image from the VLT. We found that participants were more accurate in identifying the outcome of the optimal scenes compare to the suboptimal scenes. That is, explicit knowledge was more accurate for the 80% win scene versus 20% win scene for the win pair and 20% loss scene versus 80% loss scene for the loss pair. This optimality effect did not differ between win scenes and loss scenes. The present results differ from previous findings demonstrating a more pronounced optimality effect for win scenes than for loss scenes (Lin et al. 2020). Since this memory task occurred after two sessions of the VLT, one possible explanation for these different results is that proactive interference led to overall lower memory performance. Nonetheless, what is consistent is a dissociation between learned value and explicit knowledge of the outcome contingencies. In the present study, while the learning asymmetry was present during both sessions, explicit knowledge for wins and losses appeared to be symmetrical, as demonstrated with performance on the memory task. The symmetry in explicit knowledge further supports the possibility that participants acquired some abstract knowledge of the task structure. Based on the results from the memory task, it is possible that participants generated some knowledge about the outcome contingencies, which led them to learn the optimal choice for each pair faster during the second session compared to the first session.

Limitations and Future Directions

The present study provides further evidence of the robust learning asymmetry in the VLT. While individuals differed in the extent to which they learned gains better than losses, or vice versa, there was no strong evidence for intraindividual consistency in the learning patterns for the VLT, or evidence that learning patterns were associated with traits such as impulsivity, or sensitivity to reward and punishment. Therefore, one alternative possibility is that the learning asymmetry may be due to the task structure. Thus, a question that remains is whether there are boundary conditions for this learning asymmetry. It is possible that altering the context in which learning of wins and loss associations occurs will affect how and what participants experience in the task, which in turn influences the learning pattern for win and loss associations. Future studies can provide insight into which aspects of the task contribute to the learning asymmetry by varying the parameters of the task (e.g., blocked versus interleaved presentation).

Another possibility is that people derive some general knowledge about the structure of the task or develop some strategies that influence their performance during the second session. When examining behavioral performance in the VLT, participants learned the optimal choice for both wins and losses with fewer trials during the second session. Based on the data from the current study, we are unable to speculate the general knowledge that participants might have acquired during the first session that may have helped their performance during the second session. For example, participants could have learned how many pairs of scenes to expect, that each scene pair is assigned to one valence condition, or that each scene has a different probability of leading to win and loss outcomes. If participants do derive some general knowledge about the task structure or develop strategies that influence their performance during the second session, this may obscure individual differences in response to rewards and punishments. Thus, future studies are needed to disentangle the effect of task experience on performance in the VLT and potential intraindividual consistency in learning about wins and losses. However, Lin et al. (2020) found that the learning asymmetry persisted despite participants receiving instructions about the outcome contingencies that could be associated with each scene prior to starting the VLT. Participants were presented with a table illustrating that there were three stimulus pairs, that members of the win/loss pairs were associated with specific outcome probabilities, along with the optimal strategy they would need to adopt for each valence. However, they were not shown the images in advance nor were they told which image pair was assigned to wins or losses, or the predictiveness of each scene. Therefore, it is unlikely that knowledge of the outcome contingencies that may have been acquired during the first session contributed to faster learning in the second session. Another possibility is that participants adopted a different strategy during the second session, which aided their performance. To explore this possibility, we conducted a paired-samples *t*-test for the learning rate and inverse temperature parameter, two parameters that characterize choice behavior. The results showed that participants had a higher learning rate during the second session when compared to the first session, $t(55) = -2.43$, $p = .018$, whereas the inverse temperature parameter did not differ, $t(55) = -1.30$, $p = .20$. Thus, in the second session, participants appeared to have relied more on feedback from recent trials to update the expected value of stimuli. Results from Session 1 provided evidence that higher learning rate yields a smaller learning difference between wins and losses. Thus, it is possible that there is some meta-learning that occurs across sessions. Future studies can further examine what knowledge participants may carry from the first session of learning to the second and further elucidate the strategies that participants may adopt during each session.

Conclusion

In conclusion, the present work demonstrated that the advantage of learning win associations versus loss associations persisted across two sessions of participants completing the VLT, although the learning asymmetry appeared less robust in the second session. While the advantage of learning of win associations appeared at the group level, individuals differed in the extent to which they learned gains better than losses, or vice versa. However, when evaluating consistency in performance on the VLT, there was no strong evidence for intraindividual consistency in learning patterns, or evidence that learning patterns were associated with traits such as impulsivity, or sensitivity to reward and punishment. Reinforcement learning parameters provided insight into some individual consistency that was not evident from an analysis of just the final level of learning in the VLT. The extent to which participants incorporated feedback from previous trials to generate value estimates was correlated across sessions, although there was no evidence for intraindividual consistency in the balance between exploration and exploitation. Altogether, these results suggest that there may be other factors that contribute to the learning asymmetry, such as the structure of the task. Therefore, future studies can test whether varying parameters of the task can affect the learning pattern for win and loss associations to elucidate which aspects of the task contribute to the learning asymmetry.

References

Aberg, K. C., Doell, K. C., & Schwartz, S. (2016). Linking individual learning styles to approach-avoidance motivational traits and computational aspects of reinforcement learning. *Plos ONE*, *11*(11), e0166675.doi:10.1371/journal.pone.0166675

Altman, D. G., & Royston, P. (2006). The cost of dichotomising continuous variables. *BMJ (Clinical research ed.), 332*(7549), 1080. https://doi.org/10.1136/bmj.332.7549.1080

Cáceres, P., & San Martín, R. (2017). Low Cognitive Impulsivity Is Associated with Better Gain and Loss Learning in a Probabilistic Decision-Making Task. *Frontiers in psychology*, *8*, 204. https://doi.org/10.3389/fpsyg.2017.00204

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, 67, 319-333.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*(6), 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B., Dolan, R.J., (2006). Cortical substrates for exploratory decisions in humans. *Nature 441*, 876–879.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160.

Frank, M. J., Woroch, B. S., & Curran, T. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron*, *47*(4), 495–501. https://doi.org/10.1016/j.neuron.2005.06.020

Hao, C., Cabrera, L., Lin, Z., Reuter-Lorenz, P. A., & Lewis, R. L. (July 2019). Explaining Valence Asymmetries in Human Learning in a Symmetric Value Learning Paradigm: A Computational Reinforcement Learning Account. Presented at the 4th Multidisciplinary Conference on Reinforcement Learning and Decision Making, Montréal, Québec, Canada.

Khdour, H. Y., Abushalbaq, O. M., Mughrabi, I. T., Imam, A. F., Gluck, M. A., Herzallah, M. M., & Moustafa, A. A. (2016). Generalized anxiety disorder and social anxiety disorder, but not panic anxiety disorder, are associated with higher sensitivity to learning from negative feedback: Behavioral and computational investigation. *Frontiers in Integrative Neuroscience, 10(20)*

Kim, S. H., Yoon, H., Kim, H., & Hamann, S. (2015). Individual differences in sensitivity to reward and punishment and neural activity during reward and avoidance learning. *Social cognitive and affective neuroscience*, *10*(9), 1219–1227. https://doi.org/10.1093/scan/nsv007

Lin, Z., Cabrera-Haro, L. E., & Reuter-Lorenz, P. A. (2020). Asymmetrical learning and memory for acquired gain versus loss associations. Cognition, *202*:104318. https://doi.org/10.1016/j.cognition.2020.104318

O'Brien, J. L., & Raymond, J. E. (2012). Learned predictiveness speeds visual processing. *Psychological Science, 23*(3), 359-363. doi: 10.1177/0956797611429800

Painter, D. R., Kritikos, A., & Raymond, J. E. (2014) Value learning modulates goal-directed actions. *The Quarterly Journal of Experimental Psychology, 67*(6), 1166-1175, doi: 10.1080/17470218.2013.848913

Palminteri, S., Kilford, E.J., Coricelli, G., & Blakemore, S. (2016). The Computational Development of Reinforcement Learning during Adolescence. *PLoS Computational Biology, 12*(6): e1004953. doi:10.1371/journal.pcbi.1004953

Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt impulsiveness scale. *Journal of Clinical Psychology, 51*, 768-774.

Raymond, J. E., & O'Brien, J. L. (2009). Selective Visual Attention and Motivation: The Consequences of Value Learning in an Attentional Blink Task. Psychological Science, 20(8), 981–988. https://doi.org/10.1111/j.1467-9280.2009.02391.x

Rissman, J., Gazzaley, A., & D'Esposito, M. (2009). The effect of non-visual working memory load on top-down modulation of visual processing. *Neuropsychologia, 47*(7), 1637-1646.

Rutherford, H. J. V., O'Brien, J. L., & Raymond, J. E. (2010). Value associations of irrelevant stimuli modify rapid visual orienting. *Psychonomic Bulletin & Review*, *17*(4), 536–542. https://doi-org.proxy.lib.umich.edu/10.3758/PBR.17.4.536

Scheres, A., & Sanfey, A. G. (2006). Individual differences in decision making: Drive and Reward Responsiveness affect strategic bargaining in economic games. *Behavioral and brain functions: BBF*, *2*, 35. https://doi.org/10.1186/1744-9081-2-35

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime User's Guide*. Pittsburgh: Psychology Software Tools Inc.

Sutton, R. S., Barto, A. G. (1998) Reinforcement learning. Cambridge, MA: MIT.

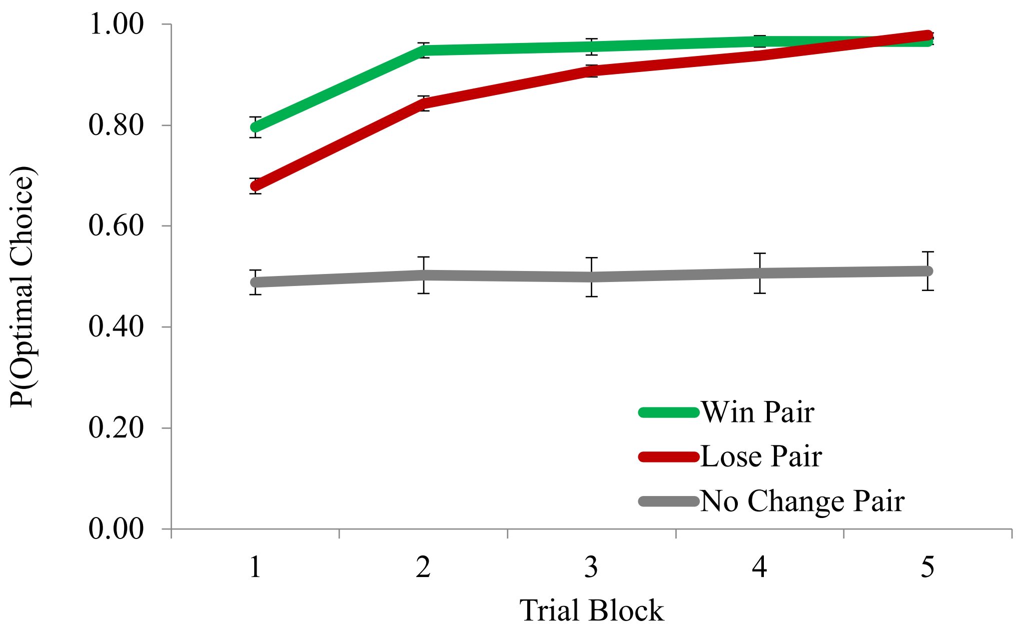
Torrubia, R., Ávila, C., Moltó, J., & Caseras, X. (2001). The sensitivity to punishment and sensitivity to reward questionnaire (SPSRQ) as a measure of Gray's anxiety and impulsivity dimensions. *Personality and Individual Differences, 31*(6), 837-862.

van den Bos, W., Cohen, M. X., Kahnt, T., & Crone, E. A. (2012). Striatum–medial prefrontal cortex connectivity predicts developmental changes in reinforcement learning. *Cerebral Cortex*, *22*(6), 1247–1255. https://doi-org.proxy.lib.umich.edu/10.1093/cercor/bhr198

Supplemental Material

We explored whether different learning patterns in the Value Learning Task (VLT) emerged for the Lin et al. (2020) data set ($N = 191$). A learning difference was calculated for each participant by subtracting the probability of optimal choice in the final block of trials, for the loss pair from the average for the win pair. We then determined the median of this learning difference (Median = 0.05) and created two groups: 1) participants with scores less than or equal to the median ($n = 92$), and 2) participants with scores greater than the median ($n = 99$). To compare learning on the VLT between these two groups, a Learner Group x Valence x Block mixed ANOVA was conducted with Learner Group as a between-subjects variable, Valence and Block as within-subjects variables, and probability of optimal choice as the dependent variable. The results revealed significant main effects for Valence, $F(1, 189) = 162.16$, $p < .001$, $\eta^2_p = .46$, Block, $F(2.41, 455.08) = 226.93$, $p < .001$, $\eta^2_p = .55$, significant two-way interactions between Learner Group x Valence, $F(1, 189) = 21.72$, $p < .001$, $\eta^2_p = .10$, Valence x Block, $F(2.61, 493.51) = 8.53$, $p < .001$, $\eta^2_p = .04$, and a significant three-way interaction between Learner Group x Valence x Block, $F(2.61, 493.51) = 7.23$, $p < .001$, $\eta^2_p = .04$. Follow-up pairwise comparisons for the three-way interaction revealed that for one group, the probability of optimal choice was significantly higher for the win pair compared to the loss pair during Blocks 1-4, ps < .02 but the probability of optimal choice was higher for the loss pair compared to the win pair during the final block, $p = .04$. In contrast, for the second group, the probability of optimal choice was significantly higher for the win pair compared to the loss pair across all blocks, ps < .001 (Figure 2.17). Thus, what differentiated these two groups is whether they learned the optimal choice for the loss pair better than the win pair, or vice versa, during the final block of trials. Therefore, these groups are referred to as "loss learners" and "win learners," respectively.

a)



b)

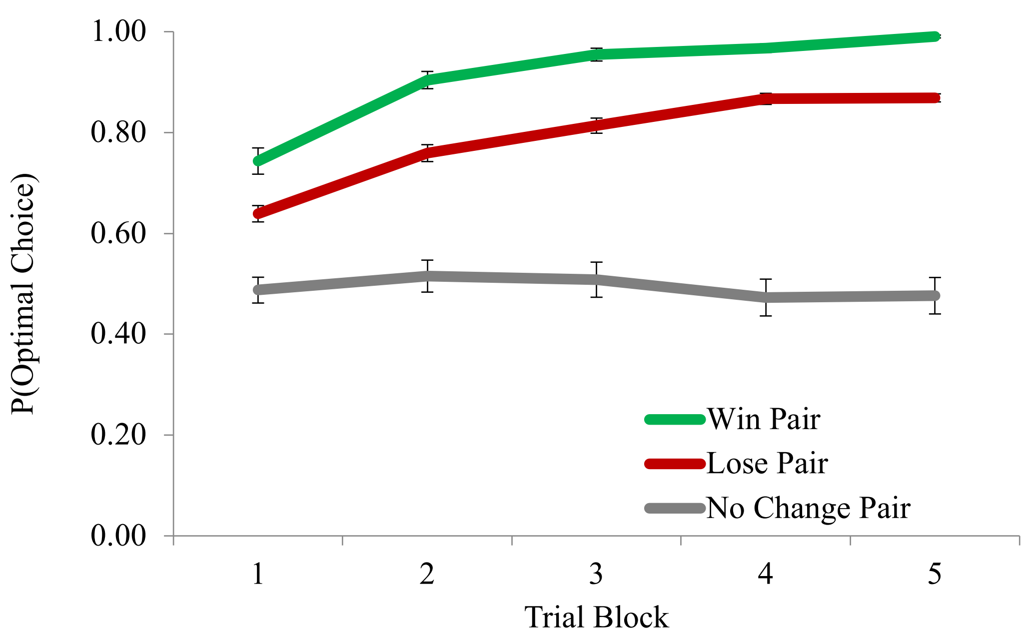


Figure 2.17. Value Learning Task Performance for Lin et al. (2020) data set.
The probability of optimal choice for participants with a learning difference score a) at or below the median level and b) above the median. The learning difference score was calculated by subtracting the probability of optimal choice in the final block of trials, for the loss pair from the average for the win pair.

Chapter 3

Separating Learning for Win and Loss Associations

Introduction

Some of the decisions we make may be influenced by the way that critical information is presented. For example, when it comes to risky decision-making, whether information is presented with a positive or negative connotation (i.e., losses or gains) can influence our willingness to take risks. People tend to avoid risk when information is presented in a positive frame but seek risk when information is presented in a negative frame, even if the absolute value of the outcomes are the same (Tversky & Kahneman, 1981). The valence (i.e., positive or negative) of information about personal attributes or behaviors can also influence our impression of individuals. Learning something negative about a new acquaintance carries more weight in forming an impression of that individual compared to learning something positive about the person (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001). The effect of positive and negative framing on people's judgments and behavior highlights our tendency to attend to valence (e.g., Baumeister et al., 2001; Rozin & Royzman, 2001). Attending to valence may be beneficial in that learning about positive and negative features in our environment or properties of stimuli can help us make decisions that will likely lead to rewards and avoid punishments.

The context in which we learn about positive and negative characteristics can take many forms. One example is the order or sequence in which information is encountered. The sequence of information has been shown to influence psychological processes in a variety of domains such as how we form impressions of other people, how we respond to surveys, and what information we remember (Asch, 1946; Gawronski, Rydell, Vervliet, & De Houwer, 2010; Schwarz, 1999). These sequence effects also apply to how we process information that varies in valence (i.e., positive or negative). For example, research on framing effects has demonstrated that encountering a frame of a given valence (i.e., positive or negative) can 'stick' in the mind and continue to influence future judgments of information, even when subsequently experiencing a frame of an opposite valence (Ledgerwood & Boydstun, 2014; Sparks & Ledgerwood, 2017). Furthermore, this research has shown that the stickiness of an initial frame can depend on the valence. An initial negative frame has a more lasting impact than an initial positive frame. That is, people's attitudes change less in response to reframing when a negatively framed object is reframed in positive terms, compared to when a positively framed object is reframed in negative terms. Thus, differences in how we respond to positive and negative information can depend on the experience with positive versus negative contingencies.

Another aspect of context that can influence decisions is whether information of different valence is presented concurrently or in separate conditions. Prior research has demonstrated that in experience-based choice tasks, whether gains and losses are presented in the same condition versus separate conditions can modulate how people respond, psychologically or physiologically, to losses versus gains (for a review see Yechiam & Hochman, 2013). According to Yechiam and Hochman (2013), losses drive how attention is invested towards a task and sensitivity towards task reinforcements. As such, whether attention is also directed towards gains will depend on whether or not gains are presented in the same condition as losses. When gains and losses are presented separately (i.e., in different conditions or temporally distant sessions), an increase in task attention is predicted only in the loss condition compared to how much attention is directed towards a task condition involving gains. This increased attention leads to greater sensitivity to payoffs in this condition (i.e., losses). As a result, an asymmetry arises such that losses modulate behavior more than gains. On the other hand, when gains and losses are presented in the same condition, the increase in task attention driven by losses, will draw attention to the overall task at hand, leading to increased sensitivity to both gain and loss reinforcements. Thus, whether gains and losses appear in the same or different conditions can affect whether an asymmetry arises in how losses and gains affect behavior. Overall, this attention framework suggests that examining the context of information may be important for predicting asymmetries in how people respond to gains and losses.

A recent study reported a learning asymmetry for gain and loss associations (Lin, Cabrera-Haro, & Reuter-Lorenz, 2020). The robustness of this asymmetry was documented in a meta-analysis of studies using a probabilistic learning task developed by Raymond and O'Brien (2009), or similar value learning task. Furthermore, in two new empirical studies, Lin and colleagues demonstrated that this learning asymmetry was evident when wins and losses led to either point or monetary earnings, and whether or not participants received explicit instructions about the outcome contingencies. A follow-up study was conducted to investigate the nature of the learning asymmetry by applying reinforcement learning theory (Hao et al., 2019). The learning asymmetry was attributed to value estimates that were more poorly discriminated for stimuli in the loss pair than stimuli in the win pair. While this modeling work provided insight into the basis for the learning asymmetry, the boundary conditions for this learning asymmetry have yet to be explored. In light of prior research documenting that the context in which positive (gains) and negative (losses) information is presented can play a role in how people perceive and act on the information, we hypothesize that context effects may influence the learning, and thus the asymmetry, for gain and loss associations. Therefore, the goal of the present study was to test this hypothesis.

In the Value Learning Task (VLT; Raymond & O'Brien, 2009), the standard presentation of win, loss and no-change trials is a random interleaved presentation requiring that people learn all associations concurrently. Thus, a question that emerged is whether this interleaved presentation selectively disadvantages loss learning. Consistent with previous research documenting that whether gains and losses are presented in the same condition versus separately can modulate asymmetric effects in gains versus losses, it is possible that the presentation format (intermixed vs. blocked) for win and loss trials may also influence the learning pattern for win and loss associations. In particular, learning loss associations may be less efficient in the context of learning concurrent win and no-change associations, than if the associations were acquired in separate blocks. Therefore, the present experiment tested the hypothesis that the presentation of win, loss, and no change trials (intermixed or blocked) will modulate the learning pattern for win and loss associations. More specifically, the learning advantage for wins compared to losses evident when win and loss trials are intermixed, will be diminished when win and loss trials are presented in separate blocks.

To test this hypothesis, we systematically varied the context in which win and loss trials appeared and measured the effects on learning. The win, loss, and no change pairs were presented in different trial blocks allowing participants to learn the outcomes associated with each pair separately. Furthermore, presenting each pair in separate blocks also allowed us to examine whether valence order (i.e., the order of win, loss, and no change blocks) affects whether a learning asymmetry arises, thus introducing another dimension of task context. Three stimulus pairs, one pair for each valence, yields six possible valence orders, which were included as a between-groups factor to explore their effects on learning of win and loss associations. Thus, different valence orders produced different prior experiential contexts for learning.

Following the VLT, we included a memory task to examine how learning context affects people's explicit memory for the outcomes associated with the images in the VLT. The results of this task may provide clues about the memory representations that are formed during the VLT and whether they vary depending on the context in which participants experienced win, loss, and no change trials.

Furthermore, we used reinforcement learning theory to develop a computational characterization of choice behavior on the VLT. A computational approach provides the benefit of examining choice behavior at a trial-level basis and captures complex learning behaviors with a few meaningful parameters. We derived the learning rate (α), which describes the extent to which new information replaces old information, and balance between exploration and exploitation (i.e., inverse temperature parameter, β) (Daw, et al., 2011; Sutton & Barto, 1998). The goal of deriving these parameters was to examine whether the learning rate and balance between exploration and exploitation varies due to the context in which win, loss, and no change trials are presented. Differences in the learning rate and inverse temperature parameters may provide insight into different learning patterns that emerge in the VLT as a function of valence order.

Method

Participants

One-hundred-ninety-six young adults (143 females; with age range 18-30 years, M = 19.80, SD = 1.90) were recruited from the University of Michigan in Ann Arbor, MI and received $10 per hour for their participation. The minimum number of participants required was determined by an *a-priori* power analysis indicating that 20 subjects were needed in each group to have 80% power for detecting a medium-sized effect when employing a $p < .05$ criterion of statistical significance (G*power; Faul, Erdfelder, Buchner, & Lang, 2009). All participants provided written informed consent for the study, which was approved by the University of Michigan, Institutional Review Board. All participants were right-handed and prescreened to exclude those with a history of depression, anxiety, ADHD, head injury, or currently taking medications that affect cognition. Following the procedure used in several prior studies of the VLT and our previous work, we adopted a learning criterion for the win and loss conditions (Lin et al. 2020, O'Brien & Raymond, 2012; Painter, Kritikos, & Raymond, 2014; Rutherford, O'Brien, & Raymond, 2010). To be included in the analysis, participants should select the optimal choice on at least 65% of the trials in the final block of the VLT for each valence pair. To ensure we had equal number of participants who reached the learning criterion for each block order, we continued data collection until achieving a total of 20 learners in each valence order condition. The final sample used for data analysis was 120 young adults (89 females, with age range 18-28 years, M = 19.63 years, SD = 1.73).[6]

Materials and Procedure

Stimuli. Eighteen grayscale images of landscape scenes were used as the visual stimuli for the VLT and Memory Task (6 were used for the VLT and Memory Task, and an additional 12 images were used as new images for the Memory Task). The images were obtained from the Cognitive Neuroscience Research Lab led by Adam Gazzaley (Rissman, Gazzaley, & D'Esposito, 2009). The size of the images were 225 pixels wide x 225 pixels tall. All stimuli

[6] The analyses reported in this chapter were repeated for the full sample and are included in the Supplemental Material.

were presented using E-prime 2.0 (Schneider, Eschman, & Zuccolotto, 2002) on a 40-cm monitor (75-Hz refresh, 1152 x 864 resolution) of a Pentium 4 computer, with a viewing distance was approximately 45 cm.

Value Learning Task. The VLT was a modified version of the task used by Raymond and O'Brien (2009). This task was previously described in Chapter 2 of this dissertation. Six scenes were selected for the VLT and divided into three pairs. Each pair was randomly assigned to one of three valence conditions: win, loss, or no change. The pairing of scenes and assignment of each pair to valence were randomized and counterbalanced across participants. The task consisted of a total of 6 blocks of trials with only one pair (win pair, loss pair, or no change pair) presented during a given block, resulting in 2 blocks per pair. Two blocks of the same valence were presented sequentially (e.g., the two blocks of the win pair were always presented one after the other). However, the order of win, loss, and no change blocks was counterbalanced across participants. Participants were randomly assigned to one of the six Valence Order conditions: Win-Lose-No Change, Lose-No Change-Win, No Change-Win-Lose, Lose-Win-No Change, Win-No Change-Lose, No Change-Lose-Win. Each pair was presented 100 times across the span of two blocks, yielding a total of 300 trials across all pairs. We opted to divide the 100 trials per pair by presenting 60 trials in the first block, and 40 trials in the second block. This was done to satisfy the following constraints: 1) we could examine choice behavior averaging across 20 trials, consistent with our previous work, and 2) to avoid averaging choice behavior across trials that included a break in between blocks. Participants were given a one-minute break between blocks. The total time on the VLT ranged between 25 to 30 minutes.

On each trial, a pair of scenes was presented on a computer screen, one above and one below a central fixation cross (Figure 3.1). Scene pair assignments were consistent across the duration of the task. The location of the image (top vs. bottom) was randomized across trials. On each trial, participants selected one scene from the pair, with the goal of maximizing point earnings. Participants pressed the “f” key to select the top scene or “j” key to select the bottom scene. For the win pair, scene selection resulted in either a win or no change. For the loss pair, scene selection resulted in either a loss or no change. For the no change pair, selecting either scene led to no change in points earned. Each image remained on the screen until participants made a selection.

Figure 3.1. Sample trial of the Value Learning Task.
A pair of neutral landscape scenes is presented on each trial, with one image above a central fixation cross and the other image below a central fixation cross. The scene pair remained on the screen until participants made a selection. Participants selected one scene from a pair by pressing one of the two designated keys on the computer keyboard. Upon selecting a scene, a yellow border appeared around the scene selected, which remained on the screen for 500 msec. The resulting outcome (win, loss, or no change) and running total of the earnings was displayed on the screen, which remained for 1500 msec. The ITI was 500 msec. In this example, a participant selected a scene that resulted in a win.

For the win and loss pairs, the probability of each outcome varied (high vs. low). One scene produced the corresponding win or loss outcome with a high probability (.80) while the other scene resulted in a win or loss outcome with a low probability (.20). The valence and the contingencies assigned to each image were constant for the entire six-block session. Participants were instructed to make the selection that would maximize points. Thus, participants had to learn which scene had a higher probability of producing a win and which scene had a lower probability of producing a loss. In other words, participants should learn the 'optimal choice' which was defined as selecting the 80% win scene for the win pair and 20% loss scene for the loss pair.

Participants received feedback regarding their selection immediately after they chose a scene. The computer screen displayed the message "Win" in green, "Lose" in red, or "No change" in black. A win or a loss outcome resulted in a gain or loss of 5 points, whereas a "No change" outcome resulted in no change in points. A running total of the earnings appeared in the top right-hand corner of the computer screen following each selection. Participants were not given any prior knowledge about the specific associations between scenes and the probability of the corresponding outcome. Instead, they were instructed to focus on maximizing their point score.

Memory Task. Participants then completed the Memory Task, which was designed to test participant's explicit knowledge of the outcome associated with each scene from the VLT. The Memory Task consisted of the six scenes shown in the VLT along with 12 new scenes. Each scene was presented individually at the center of the computer screen until participants responded (Figure 3.2). The scenes from the VLT were presented four times each, whereas each new scene was presented twice. The number of times that each image was presented was chosen so that there was a 50% probability that an item would be old versus new. The old and new scenes were presented in a randomized order for each participant. For each scene, participants indicated the outcome the scene was most likely to be associated with (Very likely to lead to a win, Occasional win, No change, Occasional loss, or Very likely to lead to a loss), or whether it was a new image. Completion of the memory task was self-paced, and no feedback was given about participants accuracy in selecting the outcome associated with each scene. The total time on the Memory Task was approximately 5 minutes.

Figure 3.2. Schematic of the Memory Task.
This Memory Task consisted of the 6 scenes from the Value Learning Task and 12 new scenes. For each scene, participants selected with what outcome the scene was most likely to be associated. Completion of the task was self-paced and no feedback was provided.

Results

Value Learning Task

Performance on the Value Learning Task is examined by calculating the probability of making the optimal choice for the win pair, loss pair, and no change pair, respectively. The optimal choice for the win pair is the scene that has a high probability of resulting in a win (80% win), for the loss pair the scene that has a low probability of resulting in a loss (20% loss), for the no change pair one scene is arbitrarily selected as the correct scene since neither scene is an optimal choice. To maintain consistency with our prior evaluation of performance on the VLT (e.g., Lin et al., 2020), the probability of optimal choice was examined for bins of 20 trials, resulting in five 20-trial blocks for each pair. Given our focus on conditions affecting the relative advantage of win versus loss, we focused on the probability of optimal choice for these pairs. Performance for the no change pair is analyzed separately. Mean probability of optimal choice for the win pair and loss pair were then compared using a 6 (Valence Order: Win-Lose-No Change, Lose-No Change-Win, No Change-Win-Lose, Lose-Win-No Change, Win-No Change-Lose, No Change-Lose-Win) x 5 (Trial Block: 1-5) x 2 (Outcome Valence: win, loss) mixed-ANOVA, with Valence Order as a between-groups factor, Outcome Valence and Trial Block as within-groups factors, and probability of optimal choice as the dependent variable. Greenhouse-Geisser corrections are applied when assumptions of sphericity are violated. The *p*-value significance cutoffs for follow-up pairwise comparisons are adjusted for multiple comparisons using a Bonferroni correction (α/number of comparisons).

A significant main effect of Trial Block emerged indicating that participants' probability of selecting the optimal scene, across valence, differed as the task progressed, $F(3.07, 349.92) = 106.85$, $p < .001$, $n_p^2 = .48$. Follow-up pairwise comparisons using a Bonferroni correction (α/number of comparisons) revealed a significant increase in the probability of optimal choice across Trial Blocks, *p*s < .001, except between Block 3 and Block 4, $p = 1.00$. A significant main effect of Outcome Valence emerged indicating that participants' probability of optimal choice was higher for the win pair ($M = .87$, $SE = .01$) compared to the loss pair ($M = .82$, $SE = .01$), $F(1, 114) = 10.97$, $p = .001$, $n_p^2 = .09$. No significant main effect of Valence Order emerged, $F(5, 114) = 2.14$, $p = .066$, $n_p^2 = .09$. However, a significant interaction between Valence and Valence Order suggested that the effect of Valence depends on the Valence Order, $F(5, 114) =$

2.82, $p = .02$, $n_{\mathrm{p}}^{2} = .11$. No other interactions were significant: Valence Order x Trial Block $F(15.35, 349.92) = 1.05$, $p = .41$, Trial Block x Outcome Valence, $F(2.93, 333.45) = 2.54$, $p = .06$, Valence Order x Trial Block x Outcome Valence, $F(14.63, 333.45) = 1.21$, $p = .27$.

Follow-up pairwise comparisons were conducted for the interaction between Valence and Valence Order. The results indicated that optimal choice performance was significantly higher for the win pair compared to the loss pair for participants who viewed the scene pairs in the order Lose-Win-No Change, $p = .001$, and No Change-Lose-Win, $p = .008$ (Figure 3.3). For the remaining Valence Orders, there was no significant difference in the probability of optimal choice between the win pair and loss pair, ps > .18. Thus, the order in which participants viewed the three pair of scenes affected how well participants learned the value associated with the win pair and loss pair after 100 trials. The two Valence Orders that led to a learning asymmetry share a similarity in that the loss pair block immediately precedes the win pair block. However, there was a third condition where participants also saw the loss pair before the win pair, Lose-No Change-Win, which did not yield a learning asymmetry ($p = .20$).

To examine whether the learning asymmetry that appeared for the valence orders Lose-Win-No Change and No Change-Lose-Win was due to better learning for the win pair, worse learning for the loss pair, or both, learning for the win pair and loss pair were compared among each valence order separately. The results revealed that the probability of optimal choice for the win pair did not significantly differ among any of the valence orders. However, the probability of optimal choice for the loss pair was significantly lower for the valence orders Lose-Win-No Change, $p = .015$, No Change-Lose-Win, $p = .003$, and Lose-No Change-Win, $p = .041$, when compared to the valence order No Change-Win-Lose. No other significant differences emerged.

These results indicate that learning of the win pair appears to be less sensitive to valence order than the loss pair. Loss learning was better when the loss block followed a win block.

For the no-change pair, the probability of selecting the optimal scene was examined by a repeated-measures ANOVA as a function of Valence Order and Trial Block. The main effect of Valence Order was not significant, $F(5, 114) = 1.78$, $p = .12$, suggesting that the probability of selecting the no change scene arbitrarily selected as the 'correct' scene did not differ between groups of participants assigned to different Valence Order conditions. A significant main effect of Trial Block emerged, $F(3.88, 442.24) = 2.78$, $p = .03$, $n_{\mathrm{p}}^2 = .02$. However, pairwise comparisons showed that none of the Trial Block Order comparisons reached statistical significance, *p*s > .08. The probability of selecting the chosen "correct" no-change scene was .50 across all participants.

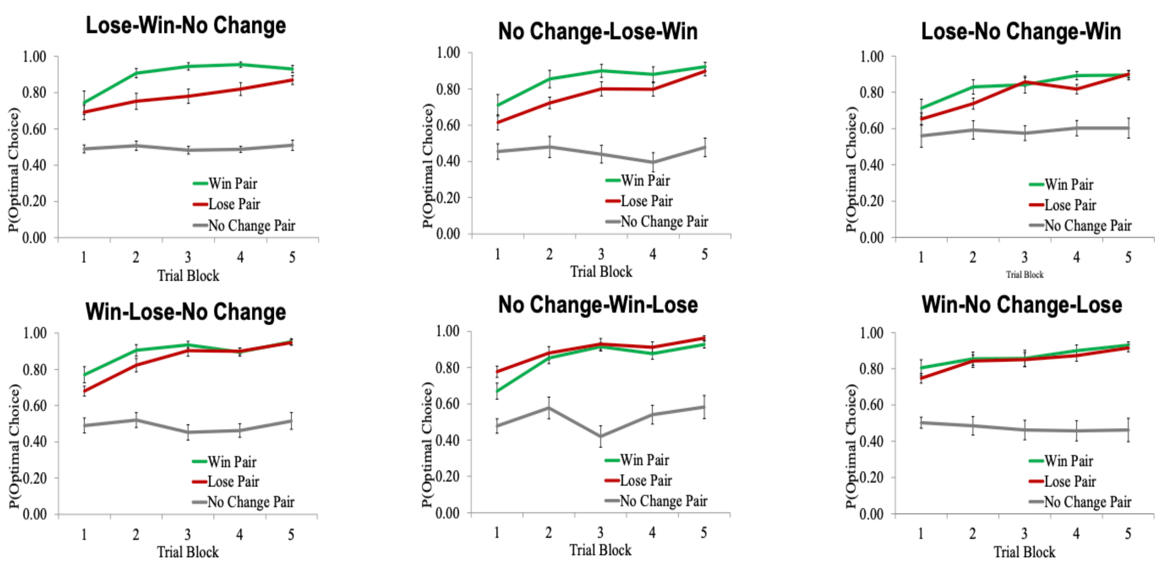


Figure 3.3. Probability of optimal choice in the Value Learning Task based on Valence Order. Probability of optimal choice was significantly higher for the win pair compared to the loss pair for the group of participants who viewed the sequence of scene pairs as Lose-Win-No Change or No Change-Lose-Win; no significant difference emerged for all remaining valence orders.

Memory Task

To examine participants' explicit knowledge of the outcomes associated with each VLT scene, a binary logistic regression analysis was performed to examine the odds of correctly selecting the outcome associated with each VLT scene (accurate outcome coded as 1, incorrect outcome coded as 0). The predictor variables were Valence Order (Win-Lose-No Change, Lose-No Change-Win, No Change-Win-Lose, Lose-Win-No Change, Win-No Change-Lose, No Change-Lose-Win, with Win-Lose-No Change arbitrarily coded as 0), Valence (win versus loss, with win coded as 0), Optimality (suboptimal versus optimal, with suboptimal coded as 0), and interaction between Valence, Optimality, and Valence Order. For the win pair, the optimal scene was defined as the 80% win scene, whereas the suboptimal scene was the 20% win scene. For the loss pair, the optimal scene was the 20% loss scene, whereas the suboptimal scene was the 80% loss scene. When determining the final model for analysis, the interaction terms were added iteratively to evaluate model fit. The final model only included terms that significantly increased model fit. Thus, the final model included Valence, Optimality, the interaction between Valence and Optimality, interaction between Valence Order and Optimality, and interaction between Valence Order and Valence as predictors of outcome accuracy.

Results of the binary logistic regression indicated that Optimality, Valence Order, the interaction between Valence and Optimality, interaction between Valence Order and Optimality, and interaction between Valence Order and Valence were significant predictors of outcome accuracy (Table 3.1). The significant effect of Optimality indicated that the odds of correctly attributing the outcome for optimal scenes were higher than for suboptimal scenes. However, the interaction between Valence and Optimality suggests that the optimality effect differed by valence (Figure 3.4). While for the win scenes memory accuracy was higher for the optimal

(80% win) scenes when compared with suboptimal (20% win) scenes, for the loss scenes, the odds of correctly attributing the outcome were higher for suboptimal (80% loss) scenes than for optimal (20% loss) scenes. The significant effect of Valence Order indicated that overall, memory accuracy was lower for the valence order No-Change-Win-Lose compared to all other orders.

Furthermore, the Valence Order x Optimality interaction suggest that the Optimality effect (higher memory accuracy for the optimal scenes versus suboptimal scenes) differed among some valence orders. In order to compare the Optimality effect among valence orders, the Win-Lose-No Change valence order was arbitrarily selected as the reference condition. Follow up analyses showed that the valence orders, Lose-No Change-Win, No Change-Win-Lose, and Win-No Change-Lose had a more pronounced optimality effect, whereas the valence order, Lose-Win-No Change had a less pronounced optimality effect.

The Valence Order x Valence interaction suggests that the effect of Valence (higher memory accuracy for win scenes versus loss scenes) differed as a function of the order in which the win, loss, and no change pairs were presented. A valence effect emerged for the orders, No Change-Win-Lose, Lose-Win-No Change, and Win-No Change-Lose. For suboptimal scenes, memory accuracy was higher for loss scenes compared to win scenes. Whereas for optimal scenes, memory accuracy was higher for win scenes compared to loss scenes.

Table 3.1.
Binary Logistic Regression Predicting Memory Accuracy from Valence, Optimality, and Valence Order.

Predictor	β	OR	95% CI for OR
Intercept	-1.00**	0.37	(0.17, 0.79)
Valence	0.57	1.76	(0.98, 3.16)
Optimality	1.03***	2.79	(1.56, 5.01)
Valence Order (L-NC-W)	0.23	1.26	(0.44, 3.64)
Valence Order (NC-W-L)	-1.62**	0.20	(0.06, 0.60)
Valence Order (L-W-NC)	-0.33	0.72	(0.25, 2.09)
Valence Order (W-NC-L)	-0.47	0.62	(0.21, 1.79)
Valence Order (NC-L-W)	0.05	1.05	(0.35, 3.14)
Valence*Optimality	-2.37***	0.09	(0.06, 0.15)
Optimality*Valence Order (L-NC-W)	1.24**	3.48	(1.62, 7.54)
Optimality*Valence Order (NC-W-L)	1.42***	4.15	(1.86, 9.40)
Optimality*Valence Order (L-W-NC)	0.87*	2.39	(1.13, 5.12)
Optimality*Valence Order (W-NC-L)	1.11**	3.04	(1.44, 6.48)
Optimality*Valence Order (NC-L-W)	0.45	1.58	(0.71, 3.53)
Valence* Valence Order (L-NC-W)	-0.11	0.90	(0.42, 1.93)
Valence*Valence Order (NC-W-L)	0.58	1.80	(0.81, 4.02)
Valence*Valence Order (L-W-NC)	0.93*	2.55	(1.20, 5.46)
Valence*Valence Order (W-NC-L)	0.17	1.19	(0.56, 2.52)
Valence*Valence Order (NC-L-W)	-0.48	0.62	(0.28, 1.40)

Note: OR = Odds Ratio.
****p* < .001, ***p* < .01, **p* < .05

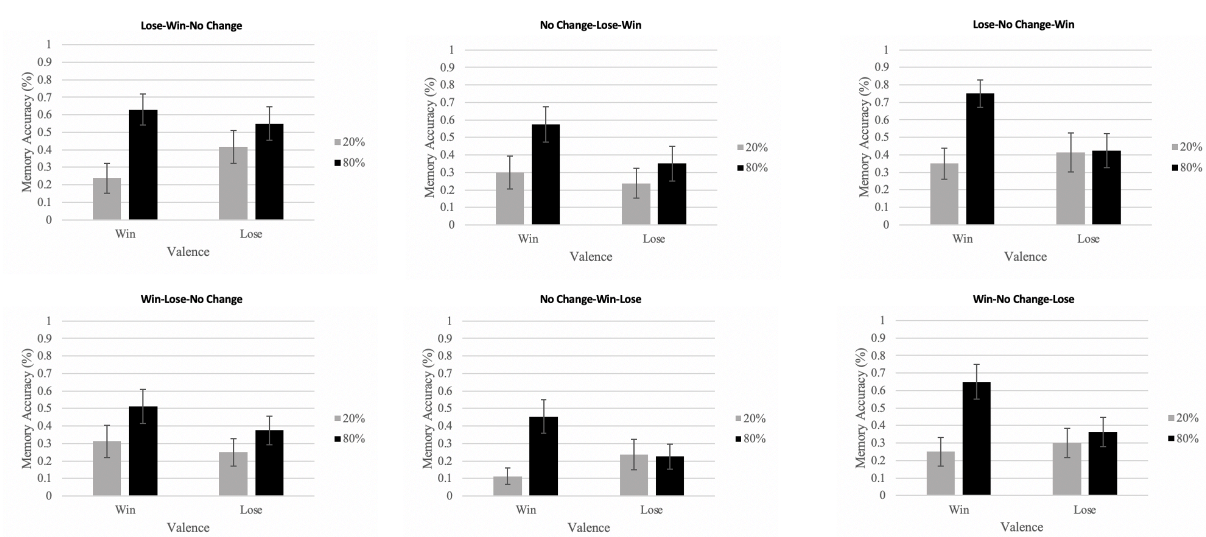


Figure 3.4. Memory data accuracy as a function of Valence and Probability of Outcomes

Outcome Attribution. We also examined participants' responses on the memory task by conducting a multinomial logistic regression with Valence, Probability, the interaction between Valence and Probability, and Valence Order as predictors of outcome response (Table 3.2). The results below are presented for each Valence Order condition.

Lose-Win-No Change. For the 80% win scene, "very likely to win" was the outcome most frequently attributed to the scene (Figure 3.5). The remaining attribution errors were distributed similarly across the remaining responses options. For the 20% win scene, "occasionally win," "no change," and "none(new)" were most frequently attributed as the outcome associated with the scene. Thus, for the win scenes, attribution errors were primarily confusing the probability of a win, valence errors were rare. Furthermore, attribution errors were more common for the 20% win scene.

For the 80% loss scene, "very likely to lose" was the outcome most frequently attributed to the scene, followed by "occasionally lose," "no change," and "none(new)" (Figure 3.6). For the 20% loss scene, "occasionally lose" was the outcome most frequently attributed to the scene. Other attributions were distributed similarly across the remaining options, except for "very likely to lose" and "none(new)" which were selected the least frequently. Thus, for the 80% loss scene, valence attribution errors were rare. For the 20% loss scene, although participants rarely attributed a high probability loss to the scene, they did make valence misattributions.

No Change-Lose-Win. For the 80% win scene, "very likely to win" was the outcome most frequently attributed to the scene. The remaining attribution errors were distributed similarly across "occasionally win," "no change," and "none(new)." For the 20% win scene, participants were just as likely to select "no change," "occasionally win," and "none(new)." Thus, for win scenes, participants rarely made a valence misattribution, but they confused the probability of a win outcome. Furthermore, attribution errors were more common for the 20% win scene.

For the 80% loss scene, "very likely to lose" was the outcome most frequently attributed to the scene. Other attributions were distributed similarly across the remaining options, except for "very likely to win" which was selected the least frequently. For the 20% loss scene, attributions were distributed similarly across all response options, with the exception of "very likely to win," which was selected the least frequently. Thus, for the loss scenes, attribution errors included confusing the probability of a loss and valence misattribution. However, participants were unlikely to misattribute a high probability win to either loss scene.

Lose-No Change-Win. For the 80% win scene, "very likely to win" was the outcome most frequently attributed to the scene. Attribution errors were rare and distributed across the remaining responses options. For the 20% win scene, "occasionally win" and "no change" were most frequently attributed as the outcome associated with the scene. Other attribution errors were rare. Thus, for the win scenes, valence misattributions were rare.

For the 80% loss scene, "very likely to lose" was the outcome most frequently attributed to the scene, followed by "occasionally lose," and "no change." For the 20% loss scene, "occasionally lose" was the outcome most frequently attributed to the scene. Other attributions were distributed similarly across the remaining options, except for "none(new)" which was selected the least frequently. Thus, for the 80% loss scene, attribution errors included confusion in the probability of a loss but rarely included valence misattributions. For the 20% loss scene, attribution errors included both confusion in the probability of a loss and valence misattribution.

Win-Lose-No Change. For the 80% win scene, "very likely to win" was the outcome most frequently attributed to the scene, followed by "occasionally win" and "none(new)." For the 20% win scene, “occasionally win” and “no change” were most frequently attributed as the outcome associated with the scene, followed by “none(new).” Thus, for the win scenes, attribution errors were primarily confusion in the probability of a win rather than a valence misattribution.

For the 80% loss scene, “very likely to lose” was the outcome most frequently attributed to the scene, followed by “occasionally lose,” “no change,” and “none(new).” For the 20% loss scene, “very likely to win” and “occasionally lose” were the outcomes most frequently attributed to the scene, followed by “occasionally win,” “no change,” and “none(new).” Thus, for the 80% loss scene, attribution errors were primarily confusion in the probability of a loss, rather than a valence misattribution. However, for the 20% loss scene, attribution errors included both valence misattribution and confusion in the probability of a loss.

No Change-Win-Lose. For the 80% win scene, “very likely to win” was the outcome most frequently attributed to the scene, followed by “occasionally win” and “none(new).” For the 20% win scene, “occasionally win,” “no change,” “occasionally lose,” and “none(new) were selected with similar frequency. Thus, for the 80% win scenes, attribution errors primarily included confusion in the probability of a win outcome. However, for the 20% win scene, misattribution errors included confusion in the probability of a win and valence misattribution.

For the 80% loss scene, outcome attributions were distributed across all possible responses, with the exception of “very likely to win,” which was selected the least frequently. For the 20% loss scene, outcome attributions were distributed across all possible responses, with the exception of “very likely to lose.” Overall, memory accuracy for the loss scenes was poor. However, participants were unlikely to misattribute a very likely win to the 80% loss scene or a high probability loss for the 20% loss scene.

Win-No Change-Lose. For the 80% win scene, “very likely to win” was the outcome most frequently attributed to the scene. The remaining attribution errors were distributed similarly across the remaining responses options. For the 20% win scene, participants were just as likely to select “no change,” “occasionally win,” “very likely to lose,” and “none(new).” “Very likely to win” and “occasionally lose” were selected the least frequently. Thus, for the 80% win scene, valence attribution errors were rare. While attribution errors for 20% win scene included valence misattribution and confusion in the probability of a win.

For the 80% loss scene, “very likely to lose” was the outcome most frequently attributed to the scene. Other attributions were distributed similarly across the remaining options, except for “none(new)” which was selected the least frequently. For the 20% loss scene, attributions were distributed across all response options, except for “none(new),” which was selected the least frequently. Thus, attribution errors for the loss scenes included valence misattribution and confusion in the probability of a loss. Overall, memory accuracy was poor for the loss scenes.

Summary of outcome attribution results. In summary, when examining the full range of responses on the memory task, several general patterns emerged. For both win scenes, the most frequent attribution errors included confusion in the probability of a win. For the 20% win scene, attribution errors also included valence misattribution. Memory accuracy was poorer for loss scenes when compared with win scenes. Attribution errors were therefore more frequent than for win scenes and included confusion in both the probability of a loss and valence misattribution. Importantly these patterns emerged across all valence orders, despite there being similar learning for win and loss associations in the VLT in some valence order conditions.

Table 3.2.
Multinomial Logistic Regression Predicting Outcome Response for each VLT scene from Valence, Probability, and Valence Order (N = 120)

Outcome Response	Valence	Probability	Valence * Probability	L-NC-W	NC-W-L	L-W-NC	W-NC-L	NC-L-W
Very likely to win	3.97***	38.32***	0.01	0.85	0.63	0.78	1.14	0.57
Occasionally win	0.70	1.83*	0.50	0.63	0.72	0.67	0.77	0.60
Occasionally lose	4.98***	1.75	0.46	1.01	0.84	1.34	1.21	0.53*
Very likely to lose	1.69*	0.97	5.37	1.01	0.45*	1.81	1.58	0.92
None (new)	0.74	3.16***	0.55	0.35***	0.62	0.64	0.47*	0.61

Note. Values in the table are odds ratio. Valence was coded as 0 for the win scenes, and 1 for the loss scenes. Probability was coded as 0 for low probability, and 1 for high probability. Win-Lose-No Change was selected as the reference condition for valence order.
*$p < .05$, ** $p < .01$, *** $p < .001$

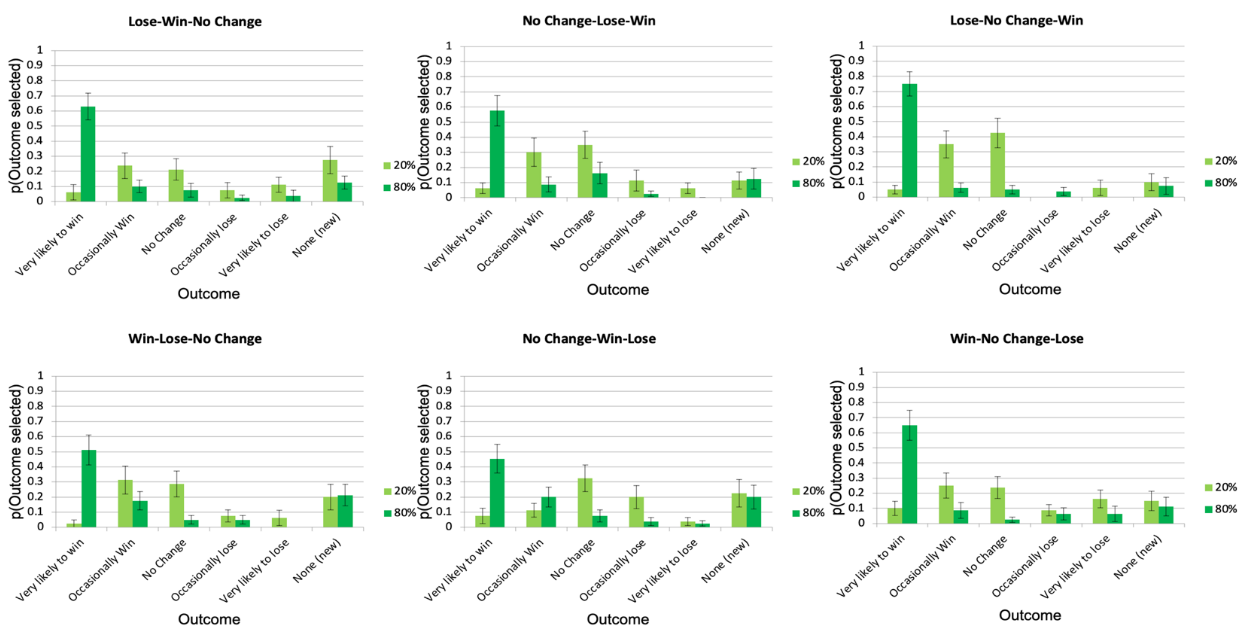


Figure 3.5. Outcome attribution responses on the Memory Task for win scenes.

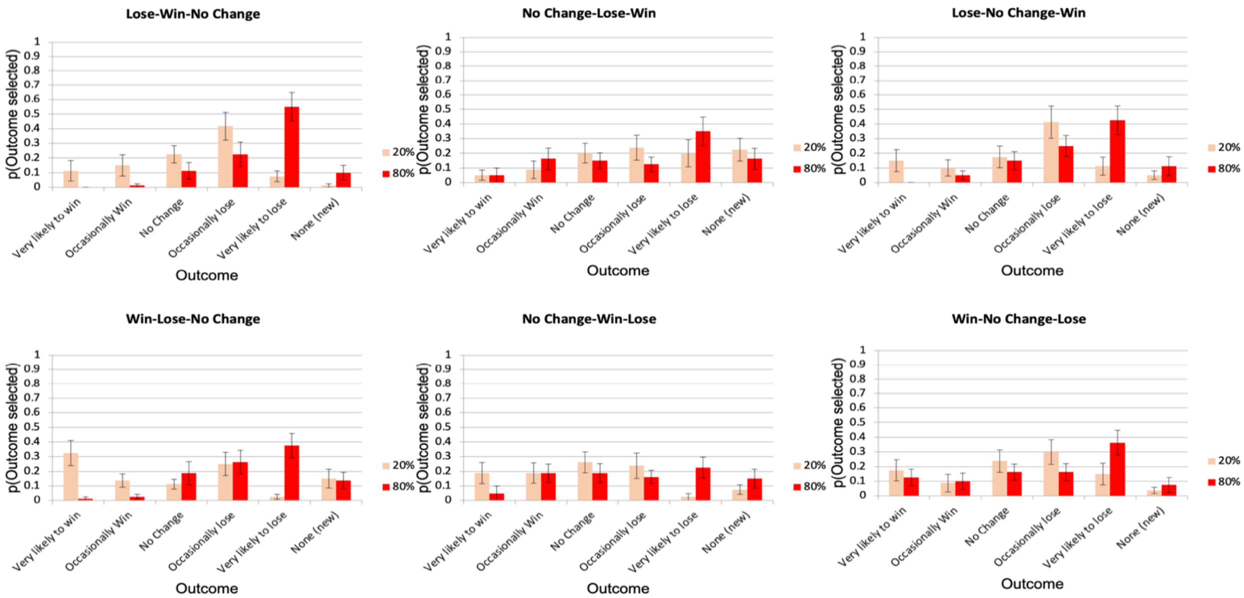


Figure 3.6. Outcome attribution responses on the Memory Task for loss scenes.

Correlation Between Value Learning and Memory Accuracy for Win and Loss Scenes

We also examined the relationship between value learning performance for win and loss scenes and accuracy for the corresponding win and loss scenes. A correlation for each image (80% win, 20% win, 80% loss, 20% loss) was computed and the significance was determined after correcting for multiple correlations using a Holm-Bonferroni correction. With the correction, no significant correlations emerged between win learning in the VLT and memory accuracy for either win scene (20%: $r(118) = -.10$, $p = .29$; /80%: $r(118) = .20$, $p = .03$). No significant correlations emerged between loss learning in the VLT and memory accuracy for either loss scene (20%: $r(118) = -.04$, $p = .64$/80%: $r(118) = -.11$, $p = .24$).

Reinforcement Learning

The learning rate and inverse temperature parameter were compared between valence order conditions using a one-way between-subjects ANOVA with Valence Order as the between-subjects variable and each parameter as the dependent variable. The results showed that the learning rate significantly differed among valence order conditions, $F(5, 114) = 4.59$, $p = .001$, $n_p^2 = .17$. Follow up pairwise comparisons indicated that the learning rate was significantly lower for participants who viewed the scene pairs in the order No Change-Lose-Win compared to participants who viewed them in the order No Change-Win-Lose, $p = .001$, and Win-No Change-Lose, $p = .028$. No other differences were significant, *p*s > .07 (Table 3.3). These results demonstrate that participants in the No Change-Lose-Win condition tended to rely more on prior knowledge more to update the expected value of the scenes compared to participants in the No-Change-Win-Lose and Win-No Change-Lose conditions. When comparing the inverse temperature parameter among valence order conditions, the results revealed that the inverse temperature parameter did not differ between valence order conditions, $F(5, 114) = 1.34$, $p = .25$.

Table 3.3.
Learning Rate and Inverse Temperature Parameter for each Valence Order Condition.

Valence Order Condition	α M(SD)	β M(SD)
Lose-Win-No Change	.22(.19)	.95(.36)
No Change-Lose-Win	.14(.13)	1.13(.87)
Lose-No Change-Win	.22(.19)	.80(.43)
Win-Lose-No Change	.22(.18)	1.02(.40)
No Change-Win-Lose	.38(.16)	.79(.33)
Win-No Change-Lose	.32(.20)	.96(.45)

Discussion

The context in which we learn about positive and negative features in our environment can affect how people respond to information, including people's attitudes and decisions (Sparks & Ledgerwood, 2017; Tversky & Kahneman, 1981). Thus, examining features of the learning context is important for understanding the differential impact of positive and negative information on choice behavior and learning. In a recent investigation of how people learn to associate neutral stimuli with win and loss outcomes, a learning asymmetry was revealed such that participants learned win associations better than loss associations (Lin et al., 2020). This learning asymmetry was reported for a value learning task in which win, loss, and no change trials are randomly interleaved, requiring win and loss associations to be learned concurrently. Thus, a question that emerged was whether this interleaved presentation selectively disadvantaged loss learning. To test the hypothesis, the present study systematically varied the context in which win and loss trials appeared. The win, loss, and no change trials were presented in different trial blocks allowing participants to learn each pair-outcome association separately. With three valenced pairs, participants were assigned to one of six valence orders: Win-Lose-No Change, Lose-No Change-Win, No Change-Win-Lose, Lose-Win-No Change, Win-No Change-Lose, or No Change-Lose-Win. We examined the effect of valence order on learning of win and loss associations, explicit knowledge of the outcomes associated with each image using a post-learning memory task, and applied reinforcement learning theory to develop a computational characterization of choice behavior.

Value Learning Task

Several key findings emerged in this task. First, participants learned to select the optimal image, leading to improved performance for win and loss trials over the course of 100 trials. Second, overall win associations were acquired more efficiently than loss associations, leading to more optimal choice performance for win than for loss trials. However, as predicted, learning context matters in that the advantage for acquiring win associations relative to loss associations, depends on valence order. Participants in the conditions Lose-Win-No Change or No-Change-

Lose-Win learned the win pair better than the loss pair. Participants who viewed the pairs in the remaining valence order conditions had similar learning for the win and loss pair.

These results show that even when separating learning for win and loss trials, a learning asymmetry still emerged in some conditions. These findings suggest that an interleaved presentation is not the sole cause of the relative disadvantage observed for loss learning in prior research (see Lin et al., 2020). However, the results from the present study demonstrate that the context in which win and loss associations are presented affects the presence or absence of the learning asymmetry. Specifically, we observed that learning for win associations was relatively stable, whereas learning for loss associations was disadvantaged only when lose learning blocks directly preceded win learning blocks.

One potential explanation for finding that the presence or absence of a learning asymmetry depends on the order of valenced blocks stems from research on framing effects. Previous research has demonstrated that encountering information of a given valence (i.e., positive or negative) can influence subsequent judgments of information of the opposite valence (Ledgerwood & Boydstu, 2014; Sparks & Ledgerwood, 2017). In the present study a learning asymmetry only emerged when loss learning directly preceded win learning and loss learning was disadvantaged rather than win learning being advantaged. One potential explanation for the present findings is that the order in which participants learn wins and losses affects the extent to which participants discount losses but has relatively no effect on participants' tendency to accept wins. Participants who first learn losses and have no prior exposure to win outcomes tend to discount the losses that result from their choice behavior. However, when prompted with the possibility of wins in a subsequent learning block, they become motivated to maximize wins in order to recover from their losses. As a result, while they may have had relatively less motivation to learn the optimal choice for the loss pair, they become motivated to learn the optimal choice for the win pair. This results in a learning asymmetry for win and loss associations, where loss learning is disadvantaged. On the other hand, participants who first learn the optimal choice for wins become focused on preserving their wins. Therefore, their motivation to avoid losses is similar to their motivation to seek wins. This results in equivalent learning of outcome associations for wins and losses. This potential explanation could be tested by increasing participants' task motivation for losses relative to wins and examining whether the learning asymmetry is reduced.

The attentional framework proposed by Yechiam and Hochman (2013) suggests that losses drive how attention is invested towards a task and in turn sensitivity towards task reinforcements. Asymmetries in how people respond to gains versus losses depends on whether gains appear in the same condition as losses. When gains and losses are presented separately, an increase in task attention is predicted only in the loss condition leading to greater sensitivity to loss payoffs versus gain payoffs. On the other hand, when gains and losses are presented in the same condition, the increase in task attention driven by losses, will draw attention to the overall task at hand, leading to increased sensitivity to both gain and loss reinforcements. Based on the results from the present study, the order in which participants experience gains and losses may also drive attention toward task reinforcements. While learning win associations was relatively stable across valence order conditions, learning loss associations was disadvantaged when loss preceded wins. However, in the present study, the temporal order of wins appears to drive motivation toward losses. That is, whether or not wins precede losses will affect whether losses are discounted. Thus, in the context of value learning, attention and sensitivity toward loss payoffs may vary as a function of whether they appear in the same condition as wins, and the temporal order of wins and losses. When wins and losses appear in separate conditions without prior exposure to a win condition, loss payoffs may receive less attention when they precede wins. This proposed explanation would suggest a revision to the predictions posed by the attentional framework to account for instances in which wins drive attention toward loss payoffs.

Memory Task

The results from the post-learning memory task showed that for the win pair, memory was better for the optimal scene (80% win scene) compared to the suboptimal scene (20% win scene). This pattern corresponds with their choice behavior in the VLT in that optimal performance required that participants select the 80% win scene. For the loss pair, the pattern was reversed: memory was better for the suboptimal scene (80% loss scene) compared to the optimal scene (20% lose scene), even though in the VLT they selected the optimal scene far more frequently than the suboptimal scene. That is, participants generally had better explicit knowledge of the outcome associated with the image that they should *select* to maximize wins and the image they should *avoid* to minimize losses. This pattern held across all valence order conditions. These results show that explicit memory differed for win- and loss-associated stimuli. Furthermore, they suggest a dissociation between learned value and explicit knowledge of outcome contingencies. While the context in which win and loss associations are presented affects the presence or absence of the learning asymmetry, memory accuracy of the outcome associated with each scene tended to be similar despite varying the context in which win and loss trials appeared. This pattern suggests that while participants may have had similar conscious strategy across valence order conditions, as reflected in memory performance, their choice behavior was influenced by context.

Reinforcement Learning

In the present study we employed a reinforcement learning framework to develop a computational characterization of choice behavior for participants in each valence order condition. We compared the learning rate, which characterizes to what extent newly acquired information overrides old information when updating the value of each scene, and the inverse temperature parameter, which describes the balance between exploration and exploitation in choice behavior. The results demonstrated that while the inverse temperature parameter did not significantly differ by valence order condition, participants in the No Change-Lose-Win condition tended to exploit prior knowledge more so to update the expected value of the scenes (i.e., lower learning rate) compared to participants in the No-Change-Win-Lose and Win-No Change-Lose conditions. Importantly, the No Change-Lose-Win condition was one valence order condition where a learning asymmetry was observed. These results suggest that the order in which participants learn win and loss associations may affect the extent to which they rely on the trial history of feedback to update the expected value of stimuli. This would coincide with previous findings that the order in which information of different valence is encountered affects how people respond to wins and losses (Yechiam & Hochman, 2013).

Limitations and Future Directions

One limitation of the current work is that the context effect for the learning asymmetry only emerged when analyzing the sample of participants who reached the learning criterion. Non-learners showed similar learning for win and losses regardless of the valence order condition. Therefore, it is possible that the context effect is contingent on reaching the learning criterion. Similar to previous studies, we adopt the learning criterion to exclude participants who show poor learning for win and loss associations (Lin et al. 2020, O'Brien & Raymond, 2012; Painter, Kritikos, & Raymond, 2014; Rutherford, O'Brien, & Raymond, 2010). Successful learning is important for studies that examine whether learned value affects subsequent processing of stimuli which has acquired value. Therefore, while adopting the learning criterion is standard, future work could explore methods to improve performance such that fewer participants are excluded.

The present study provides evidence that the presence or absence of a learning asymmetry for win and loss associations depends on the context in which win and loss trials are presented. Based on these results, we suggest that the order of valenced blocks affects the extent to which participants are motivated to discount losses and maximize wins. That is, prior experience with winning increases motivation to avoid losses, thereby leading to better learning of the outcome probabilities associated with loss images. Future work could test whether increasing the salience of losses could increase motivation to avoid losses, and thus reduce the learning asymmetry. One approach would be to introduce a loss framing. Participants could start off with an endowment, which will decrease with each loss outcome. Another approach may be to increase the magnitude of the losses. Previous work has suggested that participants are more loss averse when the magnitude of outcomes are large (e.g., Harinck,van Dijk, Van Beest, & Mersmann, 2007). Thus, it is possible that increasing the magnitude of the outcomes will motivate participants to avoid losses. Future work could test this possibility by assessing task motivation after each learning block with a self-report measure. If introducing a loss framing or increasing the magnitude of outcomes increase motivation to avoid losses, then we predict that the learning asymmetry in Lose-Win-No Change or No-Change-Lose-Win will be reduced.

While several patterns emerged from the memory task, performance on the task was relatively poor. Thus, the results could be underestimating participants' knowledge of the outcomes associated with each scene. Future work could consider revising the memory task to better estimate participants' explicit knowledge. This possibility will be discussed further in the general discussion of this dissertation.

In conclusion, the present work demonstrated that the context in which learning takes place can affect the rate of learning to associated valenced outcomes with otherwise neutral stimuli. When a learning asymmetry emerged, loss learning was disadvantaged relative to win learning, a pattern reported in our previous work. On the other hand, win learning appears to be relatively robust across variations in the learning context. We also demonstrated that whether or not a learning asymmetry is evident in the VLT, explicit knowledge of outcome valence and predictiveness differs for win and loss associations. Thus, taking into account the order in which we receive information can help us understand whether the sequence of information can play a role in gain and loss asymmetries. Loss learning appears to be vulnerable to context, especially when prior experience with wins is absent. Underlying asymmetries that favor win associations has implications for how we perceive or act on information that vary in valence and the extent to which positive information has a larger effect on subsequent behavior than losses. Findings that stimuli previously associated with wins receive preferential processing in subsequent encounters may be altered based on the conditions in which learning takes place.

References

Asch, S. E. (1946). Forming impressions of personality. *The Journal of Abnormal and Social Psychology, 41*(3), 258–290. https://doi.org/10.1037/h0055756

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5,* 323–370. http://dx.doi.org/10.1037/1089-2680.5.4.323

Gawronski, B., Rydell, R. J., Vervliet, B., & De Houwer, J. (2010). Generalization versus contextualization in automatic evaluation. *Journal of Experimental Psychology: General*, *139*(4), 683–701. https://doi-org.proxy.lib.umich.edu/10.1037/a0020315

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*(6), 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160.

Hao, C., Cabrera, L., Lin, Z., Reuter-Lorenz, P. A., & Lewis, R. L. (July 2019). Explaining Valence Asymmetries in Human Learning in a Symmetric Value Learning Paradigm: A Computational Reinforcement Learning Account. Presented at the 4th Multidisciplinary Conference on Reinforcement Learning and Decision Making, Montréal, Québec, Canada.

Harinck, F., Van Dijk, E., Van Beest, I., & Mersmann, P. (2007). When gains loom larger than losses: Reversed loss aversion for small amounts of money. *Psychological Science*, *18*(12), 1099–1105. https://doi-org.proxy.lib.umich.edu/10.1111/j.1467-9280.2007.02031.x

Ledgerwood, A., & Boydstun, A. E. (2014). Sticky prospects: loss frames are cognitively stickier than gain frames. *Journal of experimental psychology. General*, *143*(1), 376–385. https://doi.org/10.1037/a0032310

Lin, Z., Cabrera-Haro, L. E., & Reuter-Lorenz, P. A. (2020). Asymmetrical learning and memory for acquired gain versus loss associations. Cognition, *202*:104318. https://doi.org/10.1016/j.cognition.2020.104318

O'Brien, J. L., & Raymond, J. E. (2012). Learned predictiveness speeds visual processing. *Psychological Science, 23*(3), 359-363. doi: 10.1177/0956797611429800

Painter, D. R., Kritikos, A., & Raymond, J. E. (2014) Value learning modulates goal-directed actions. *The Quarterly Journal of Experimental Psychology, 67*(6), 1166-1175, doi: 10.1080/17470218.2013.848913

Raymond, J. E., & O'Brien, J. L. (2009). Selective Visual Attention and Motivation: The Consequences of Value Learning in an Attentional Blink Task. Psychological Science, 20(8), 981–988. https://doi.org/10.1111/j.1467-9280.2009.02391.x

Rissman, J., Gazzaley, A., & D'Esposito, M. (2009). The effect of non-visual working memory load on top-down modulation of visual processing. *Neuropsychologia, 47*(7), 1637-1646.

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review, 5,* 296–320. http://dx.doi.org/10.1207/S15327957PSPR0504_2

Rutherford, H. J. V., O'Brien, J. L., & Raymond, J. E. (2010). Value associations of irrelevant stimuli modify rapid visual orienting. *Psychonomic Bulletin & Review*, *17*(4), 536–542. https://doi-org.proxy.lib.umich.edu/10.3758/PBR.17.4.536

Schwarz, N. (1999). Self-reports—How the questions shape the answers. *American Psychologist, 54,* 93–105. http://dx.doi.org/10.1037/0003- 066X.54.2.93

Sparks, J., & Ledgerwood, A. (2017). When good is stickier than bad: Understanding gain/loss asymmetries in sequential framing effects. *Journal of experimental psychology. General*, *146*(8), 1086–1105. https://doi.org/10.1037/xge0000311

Sutton, R. S., Barto, A. G. (1998) Reinforcement learning. Cambridge, MA: MIT.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458. https://doi.org.proxy.lib.umich.edu/10.1126/science.7455683

Yechiam, E., & Hochman, G. (2013). Losses as modulators of attention: Review and analysis of the unique effects of losses over gains. *Psychological Bulletin, 139*(2), 497–518. https://doi.org/10.1037/a0029383

Supplemental Material

Performance on the Value Learning Task and Memory task were evaluated using the full sample of 196 participants. Furthermore, the reinforcement learning parameters were derived for all participants and compared among valence order conditions. The analyses reported below are the same as the analyses conducted for participants who reached the learning criterion.

Value Learning Task

Similar to the analysis conducted with the sample of participants who reached the learning criterion, a 6 (Valence Order: Win-Lose-No Change, Lose-No Change-Win, No Change-Win-Lose, Lose-Win-No Change, Win-No Change-Lose, No Change-Lose-Win) x 5 (Trial Block: 1-5) x 2 (Outcome Valence: win, loss) mixed-ANOVA was conducted, with Valence Order as a between-groups factor, Trial Block, and Outcome Valence as within-groups factors, and probability of optimal choice as the dependent variable. When evaluating main effects and interactions, Greenhouse-Geisser corrections were applied when assumptions of sphericity were violated. The *p*-value significance cutoffs for follow-up pairwise comparisons were adjusted using a Bonferroni correction (α/number of comparisons).

Similar to the main analysis, a significant main effect of Trial Block emerged, indicating that probability of optimal choice differed across trial blocks, $F(2.88, 546.77) = 71.20$, $p < .001$, $n_p^2 = .27$. Follow-up pairwise comparisons revealed a significant increase in the probability of optimal choice between Block 1 and 2, $p < .001$, and Block 2 and 3, $p = .003$. There was no significant increase between Block 3 and 4, $p = 1.00$, or between Block 4 and 5, $p = .11$. However, a significant interaction emerged between Valence Order and Trial Block suggesting that the main effect of Trial Block differed by Valence Order, $F(14.39, 546.77) = 1.95$, $p = .02$, $n_p^2 = .05$ (Figure 3.7). Follow-up pairwise comparisons revealed that for the Valence Order conditions Lose-Win-No Change, No Change, Lose-Win, Lose-No Change-Win, Win-Lose-No Change, and No Change-Win-Lose, there was a significant increase in optimal choice only between Trial Block 1 and 2, $p < .01$. No significant differences emerged among the remaining Trial Blocks, *p*s > .15. For the Win-No Change-Lose condition, there was no significant difference in optimal choice among any Trial Blocks, *p*s > .86.

No significant main effect emerged when evaluating Outcome Valence, $F(1, 190) = 1.45$, $p = .23$, and Valence Order, $F(1, 190) = 1.26$, $p = .28$. No significant interactions emerged when evaluating Outcome Valence x Valence Order, $F(5, 190) = 1.29$, $p = .27$, Outcome Valence x Trial Block, $F(2.98, 566.79) = 2.03$, $p = .11$, and Outcome Valence x Trial Block x Valence Order, $F(14.92, 566.79) = 1.28$, $p = .21$. Thus, when including all participants in the analysis, the learning asymmetry that emerged for the Valence Order conditions Lose-Win-No Change and No Change-Lose-Win no longer appeared.

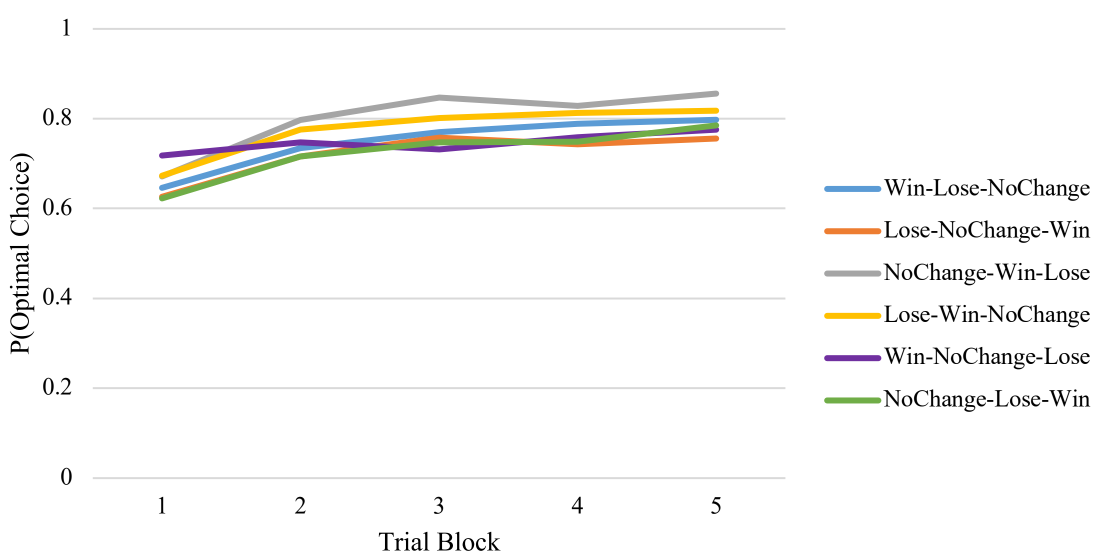


Figure 3.7. Probability of optimal choice for each Valence Order condition collapsed across valence.

The probability of selecting the arbitrarily chosen 'optimal' scene was also examined for the No Change pair by conducting a repeated-measures ANOVA with Valence Order as a between-subjects variable and Trial Block as within-subjects factor. The main effect of Valence Order was significant, $F(5, 190) = 3.45$, $p = .005$, $n_{\mathrm{p}}^2 = .08$, suggesting that the probability of selecting the optimal scene differed among groups of participants assigned to different Valence Order conditions. Follow-up pairwise comparisons using a Bonferroni correction revealed that optimal scene selection was higher for the group of participants in the Lose-No Change-Win condition ($M = .59$, $SE = 0.03$) versus No Change-Lose-Win condition ($M = .44$, $SE = 0.03$). No other significant differences emerged, ps > .07. No significant effect emerged for Trial Block, $F(3.83, 727.22) = 1.71$, $p = .15$, or the interaction between Trial Block and Valence order, $F(19.14, 727.22) = .66$, $p = .87$.

Memory Task

A binary logistic regression was conducted to examine the odds of correctly selecting the outcome associated with each VLT scene (accurate outcome coded as 1, incorrect outcome coded as 0). The predictor variables were Valence Order (Win-Lose-No Change, Lose-No Change-Win, No Change-Win-Lose, Lose-Win-No Change, Win-No Change-Lose, No Change-Lose-Win, with Win-Lose-No Change arbitrarily coded as 0), Valence (win versus loss, with win coded as 0), and Optimality (suboptimal versus optimal, with suboptimal coded as 0). Interaction terms were added to the model iteratively to examine whether there was a significant improvement in model fit. The final model only incorporated interaction terms that significantly increased model fit. Therefore, the final model included Valence, Optimality, Valence Order, and the interaction between Valence and Optimality as predictors of outcome accuracy.

Results of the binary logistic regression indicated that Valence, Optimality, and the interaction between Valence and Optimality were significant predictors of outcome accuracy (Table 3.4). The significant effect of Valence indicated that the odds of correctly attributing the outcome were higher for loss scenes than for win scenes. However, the interaction between Valence and Optimality suggests that this pattern differed by optimal versus suboptimal scenes. While for the suboptimal scenes memory accuracy was higher for losses (80% loss scenes) versus wins (20% win scenes), for optimal scenes the pattern was reversed (Figure 3.8). For optimal scenes, memory accuracy was higher for the win scenes (80% win scene) versus loss scenes (20% loss scene). The analysis with participants who reached the learning criterion also showed these results but only for the valence orders, No Change-Win-Lose, Lose-Win-No Change, and Win-No Change-Lose.

The significant effect of Optimality indicated that the odds of correctly attributing the outcome for optimal scenes were higher than for suboptimal scenes. However, the significant interaction between Valence and Optimality suggests that this pattern differed by valence. While for the win scenes memory accuracy was higher for the optimal (80% win) scenes when compared with suboptimal (20% win) scenes, the pattern was reversed for loss scenes. For the loss scenes, the odds of correctly attributing the outcome were higher for suboptimal (80% loss) scenes than for optimal (20% loss) scenes. These results are similar to the results with the sample of participants who reached the learning criterion.

Table 3.4.
Binary Logistic Regression Predicting Memory Accuracy from Valence, Optimality, and Valence Order.

Predictor	β	OR	95% CI for OR
Intercept	-1.46***	0.23	(-1.97, -0.97)
Valence	0.45***	1.57	(0.20, 0.70)
Optimality	1.28***	3.60	(1.04, 1.53)
Valence Order (L-NC-W)	0.33	1.39	(-0.30, 0.97)
Valence Order (NC-W-L)	-0.34	0.71	(-1.03, 0.34)
Valence Order (L-W-NC)	0.53	1.70	(-0.15, 1.21)
Valence Order (W-NC-L)	-0.02	0.98	(-0.66, 0.63)
Valence Order (NC-L-W)	0.03	1.03	(-0.64, 0.70)
Valence*Optimality	-1.49***	0.22	(-1.84, -1.15)

Note: OR = Odds Ratio.
****p* < .001, ***p* < .01, **p* < .05

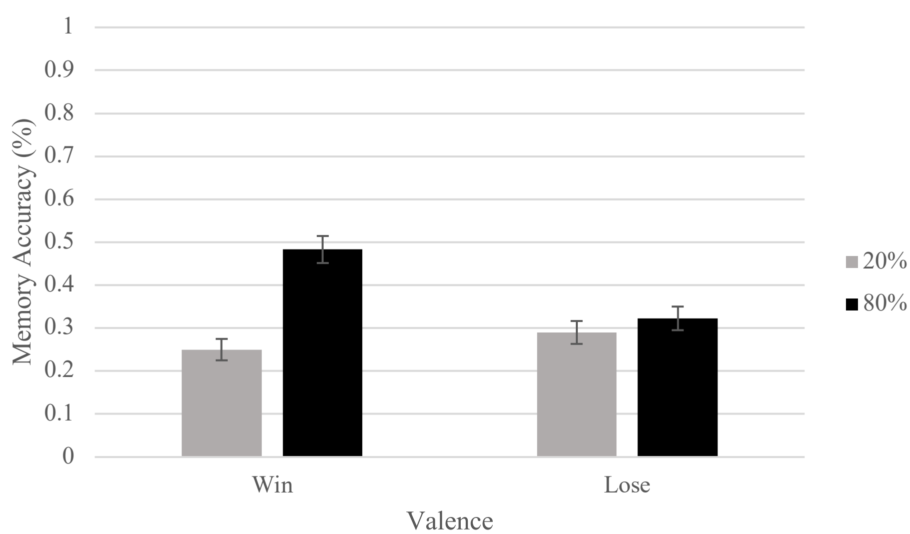


Figure 3.8. Memory data accuracy as a function of Valence and Probability of Outcome across all valence order conditions.

Outcome Attribution. As in the main analysis, we examined participants' responses on the memory task by conducting a multinomial logistic regression with Valence, Probability, the interaction between Valence and Probability, and Valence Order as predictors of outcome response (Table 3.5). The results for each Valence Order condition are presented below.

Lose-Win-No Change. For the 80% win scene, “very likely to win” was the outcome most frequently attributed to the scene, followed by “occasionally win” (Figure 3.9). Other attribution errors were distributed similarly across the remaining responses options. For the 20% win scene, “occasionally win” and “none(new)” were most frequently attributed as the outcome associated with the scene. Other attribution errors were distributed similarly across the remaining responses options. Thus, for the win scenes, attribution errors were primarily confusion in the probability of a win, while valence errors were rare. Furthermore, attribution errors were more common for the 20% win scene.

For the 80% loss scene, “very likely to lose” and “occasionally lose” were most frequently attributed to the scene (Figure 3.10). For the 20% loss scene, “occasionally lose” was the outcome most frequently attributed to the scene. Other attributions were distributed similarly across the remaining options, except for “none(new)” which was selected the least frequently. Thus, for the 80% loss scene, attribution errors were primarily confusion of the probability of a loss. However, for the 20% loss scene, attribution errors included both confusion of the probability of a loss and valence misattribution.

No Change-Lose-Win. For the 80% win scene, “very likely to win” was the outcome most frequently attributed to the scene. The remaining attribution errors were distributed similarly across “occasionally win,” “no change,” and “none(new).” For the 20% win scene, participants were just as likely to select “no change,” and “occasionally win.” Other attributions were distributed similarly across the remaining response options, except for “very likely to lose,” except for “very likely to lose,” which was selected the least frequently. Overall, for the 80% win scene, attribution errors were primarily confusion of the probability of a loss. However, for the

20% win scene, attribution errors included both confusion of the probability of a win and valence misattribution.

For the 80% loss scene, “very likely to lose” was the outcome most frequently attributed to the scene. Other attributions were distributed similarly across the remaining options, except for “very likely to win” which was selected the least frequently. For the 20% loss scene, attributions were distributed similarly across all response options, with the exception of “very likely to win,” which was selected the least frequently. Thus, for the loss scenes, attribution errors included confusing the probability of a loss and valence misattribution. However, participants were unlikely to misattribute the scenes to a high probability win.

Lose-No Change-Win. For the 80% win scene, “very likely to win” was the outcome most frequently attributed to the scene. The remaining outcome attribution were primarily distributed between “occasionally win” and “none(new).” For the 20% win scene, “occasionally win” and “no change” were most frequently attributed as the outcome associated with the scene. The remaining outcome attributions were primarily “none(new).” Thus, for the win scenes, valence misattributions were primarily confusion in the probability of a win outcome.

For the 80% loss scene, “very likely to lose” was the outcome most frequently attributed to the scene, followed by “occasionally lose,” and “no change.” For the 20% loss scene, “occasionally lose” and “no change” were most frequently attributed to the scene. Other attributions were distributed similarly across the remaining options. Thus, for the 80% loss scene, attribution errors included confusion in the probability of a loss but rarely included valence misattributions. For the 20% loss scene, attribution errors included valence misattribution and confusion in the probability of a loss.

Win-Lose-No Change. For the 80% win scene, “very likely to win” was the outcome most frequently attributed to the scene. The remaining attribution errors were primarily between “occasionally win” and “none(new).” For the 20% win scene, “occasionally win” was the outcome most frequently attributed to the scene, followed by “no change” and “none(new).” Thus, for the win scenes, attribution errors were primarily confusion in the probability of a win rather than a valence misattribution.

For the 80% loss scene, “very likely to lose” and “occasionally lose” was the outcome most frequently attributed to the scene. The remaining outcome attribution were distributed similarly between “occasionally win,” “no change,” and “none(new).” For the 20% loss scene, “occasionally lose,” “very likely to win” and “occasionally win” were the outcomes most frequently attributed to the scene, followed by “none(new),” and “no change.” Thus, for the 80% loss scene, attribution errors were primarily confusion in the probability of a loss, rather than a valence misattribution. However, for the 20% loss scene, attribution errors included both valence misattribution and confusion in the probability of a loss.

No Change-Win-Lose. For the 80% win scene, “very likely to win” was the outcome most frequently attributed to the scene, followed by “occasionally win” and “none(new).” For the 20% win scene, “occasionally win,” “no change,” “occasionally lose,” and “none(new) were selected with similar frequency. Thus, for the 80% win scenes, attribution errors primarily included confusion in the probability of a win outcome. However, for the 20% win scene, misattribution errors included confusion in the probability of a win and valence misattribution.

For the 80% loss scene, attribution errors were distributed across all possible responses, with the exception of “very likely to win,” which was selected the least frequently. For the 20% loss scene, attribution errors were distributed across all possible responses, with the exception of “very likely to lose” and “none(new).” Overall, memory accuracy for the loss scenes was poor. However, participants were unlikely to misattribute a high probability win to the 80% loss scene or a high probability loss for the 20% loss scene.

Win-No Change-Lose. For the 80% win scene, “very likely to win” was the outcome most frequently attributed to the scene. The remaining attribution errors were primarily between “occasionally win,” “no change,” and “none(new).” For the 20% win scene, responses were distributed similarly across all options. Thus, for the 80% win scene attribution errors were primarily confusion in the probability of a win. For the 20% win scene, participants had poor explicit knowledge of the outcome associated with the scene.

For the 80% loss scene, “very likely to lose” was the outcome most frequently attributed to the scene. Other attributions were distributed similarly across the remaining options, except for “none(new)” which was selected the least frequently, followed by “occasionally lose” and “no change.” For the 20% loss scene, “occasionally lose” and “no change” were most frequently selected. Other attributions were distributed across the remaining response options. Thus, attribution errors for the loss scenes included valence misattribution and confusion in the probability of a loss. Overall, memory accuracy was poor for the loss scenes.

Summary of outcome attribution results. In summary, several general patterns emerged across all valence order conditions, which were similar to patterns that emerged for participants who met the learning criterion. For both win scenes, the most frequent attribution errors included confusion in the probability of a win. For the 20% win scene, attribution errors included confusion in the probability of a win and valence misattribution. Overall, memory accuracy was poorer for loss scenes when compared with win scenes. For loss scenes, attribution errors included confusion in the probability of a loss and valence misattribution.

Table 3.5.
Multinomial Logistic Regression Predicting Outcome Response for each VLT scene from Valence, Probability, and Valence Order (N = 196)

Outcome Response	Valence	Probability	Valence * Probability	L-NC-W	NC-W-L	L-W-NC	W-NC-L	NC-L-W
Very likely to win	1.91***	12.90***	0.03***	0.39***	0.45***	0.58*	0.62*	0.41***
Occasionally win	0.76	1.76*	0.59*	0.38***	0.58*	0.71	0.43***	0.46***
Occasionally lose	3.38***	1.49	0.66	0.42***	0.67	1.20	0.65	0.38***
Very likely to lose	1.70**	0.51	7.58***	0.70	0.49*	1.61	1.06	0.79
None (new)	0.73	2.21***	0.72	0.34***	0.61*	0.79	0.57*	0.54**

Note. Values in the table are odds ratio. Valence was coded as 0 for the win scenes, and 1 for the loss scenes. Probability was coded as 0 for low probability, and 1 for high probability. Win-Lose-No Change was selected as the reference condition for valence order.
p* < .05, ** *p* < .01, * *p* < .001

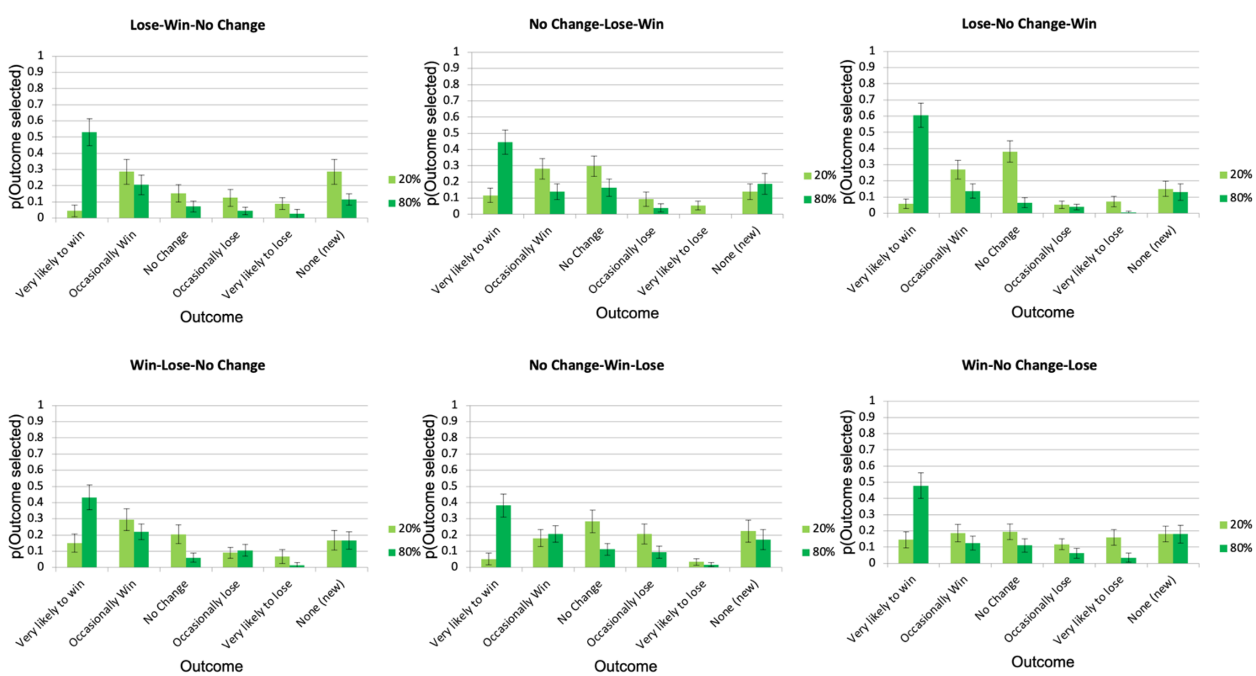


Figure 3.9. Outcome attribution responses on the Memory Task for win scenes.

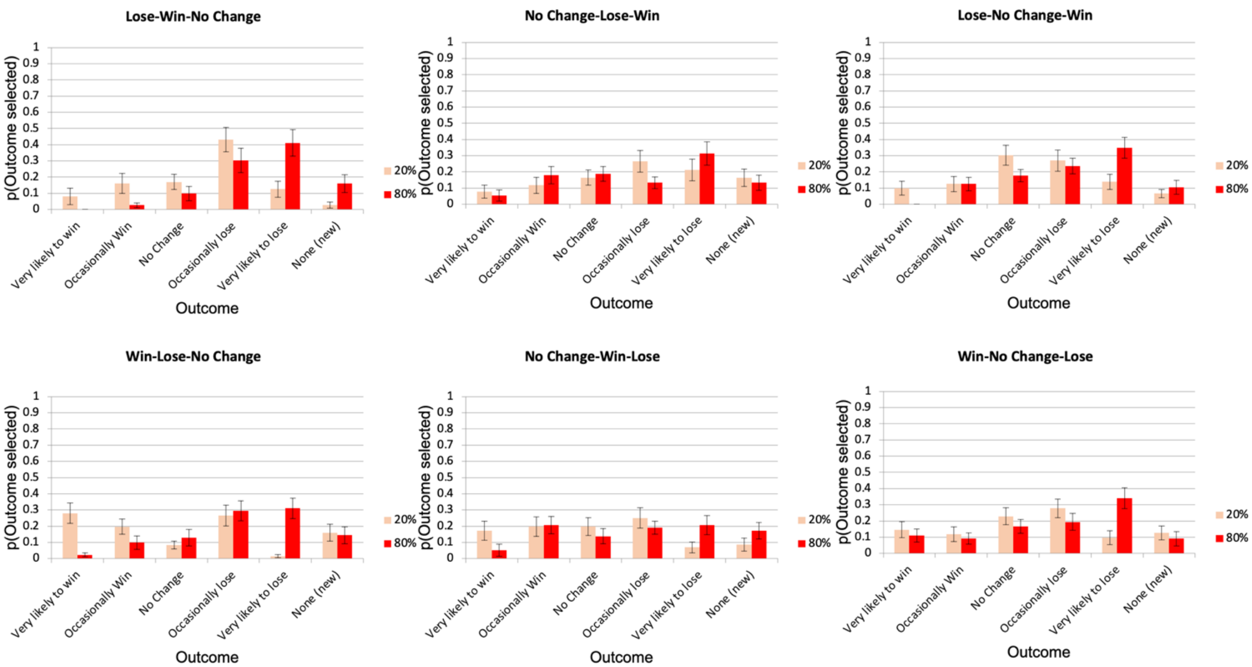


Figure 3.10. Outcome attribution responses on the Memory Task for loss scenes.

Reinforcement Learning

Similar to the analysis reported with participants who reached the learning criterion, the learning rate and inverse temperature parameter were compared between valence order conditions using a one-way between-subjects ANOVA with Valence Order as the between-subjects variable and each parameter as the dependent variable. The results revealed that the learning rate significantly differed among valence order conditions, $F(5, 190) = 2.67$, $p = .02$, $n_{\text{p}}^2 = .07$. . Follow up pairwise comparisons using a Bonferroni corrected revealed a significantly lower learning rate for participants who viewed the scene pairs in the order No Change-Lose-Win compared to participants who viewed them in the order No Change-Win-Lose, $p = .02$. This significant difference was similar to results with learners. No other differences were significant, $ps > .13$ (Table 3.6).

When comparing the inverse temperature parameter among valence order conditions, the results revealed no significant difference among valence order conditions, $F(5, 190) = .31$, $p = .90$. This non-significant difference mirrored the results from the sample of participants who reached the learning criterion.

Table 3.6.
Learning Rate and Inverse Temperature Parameter for each Valence Order Condition.

Valence Order Condition	α M(SD)	β M(SD)
Lose-Win-No Change	.23(.20)	.81(.25)
No Change-Lose-Win	.18(.20)	.76(.84)
Lose-No Change-Win	.22(.26)	.78(1.02)
Win-Lose-No Change	.22(.22)	.89(.83)
No Change-Win-Lose	.37(.22)	.70(.40)
Win-No Change-Lose	.28(.25)	.70(.52)

Chapter 4

The Effect of Outcome Magnitude on Win and Loss Associations

Introduction

Many studies in psychology have documented asymmetries in the way people respond to undesirable, harmful, or unpleasant outcomes and consequences compared to those that are desirable, beneficial, or pleasant (for reviews, see Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Rozin & Royzman, 2001; Vaish, Grossmann, & Woodward, 2008). One of the more well-known observations in decision-making research is that people are more sensitive to the possibility of losing objects or money than they are to the possibility of gaining the same objects or amounts of money, a behavior referred to as *loss aversion* (Kahneman & Tversky, 1979). That is, you are likely to be more upset about losing $50 than you are happy about gaining $50. The finding that loss outcomes have larger subjective weight than equivalent gain outcomes has been used to explain a wide range of behaviors. One example is the endowment effect which refers to the finding that people are more likely to retain an object that they already own than to endeavor to acquire that same object when they do not own it (*endowment effect*; Kahneman, Knetsch, & Thaler, 1990). This effect be explained by loss aversion in that the negative feelings associated with losing an item are stronger than the positive feelings associated with obtaining the same item.

Based on decision-making research highlighting loss aversion, one might expect that in learning to associate a stimulus with an outcome, people would be more motivated to avoid losses versus receiving rewarding outcomes and therefore prioritize learning loss outcome associations. However, recently the opposite asymmetry was identified in a value learning task (Lin, Cabrera-Haro, & Reuter-Lorenz, 2020). Through trial-and-error participants in this task learn to associate neutral stimuli (e.g., black and white landscape scenes) with probabilistic gains or losses. On each trial, their objective is to select the scene that maximizes gains or minimizes losses. Both meta-analytic evidence and a series of new experiments document that people learned stimulus-outcome associations better for images that lead to wins when compared to losses (Lin et al., 2020). This asymmetry runs contrary to what would be expected based on the observation that individuals tend to be loss averse. Therefore, the question arises as to whether the parameters typically used in the value learning task affect the presence or absence of this learning asymmetry. This question is addressed in the present research.

Boundary conditions for valence-based asymmetries. Although loss aversion is a robust phenomenon, recent work has demonstrated conditions under which this asymmetric response to losses versus gains is either diminished or reversed (e.g., Ert & Erev, 2013). One factor that influences how people may respond differently to gains and losses is the magnitude of outcomes (Erev, Ert, & Yechiam, 2008; Harinck,van Dijk, Van Beest, & Mersmann, 2007). Studies have demonstrated an absence of loss aversion when small-to-moderate amounts of money are at stake rather than large amounts of money. These effects have emerged in tasks that involve rating the pleasantness/unpleasantness of losing or gaining different amounts of money and hypothetical gambles where participants indicate how much money they would be willing to risk in order to participate in a coin-toss gamble (Harinck et al. 2007).

Harinck et al. (2007) proposed an explanation for the effect of outcome magnitude on the presence or absence of loss aversion based on the hedonic principle and psychological strategy of discounting. The hedonic principle suggests that individuals are motivated to seek and enjoy positive outcomes and avoid or discount negative outcomes (Kermer, Driver-Linn, Wilson, & Gilbert, 2006). One approach to minimizing the negative feelings associated with negative outcomes is to trivialize the loss. Since in everyday life, we are more likely to experience losses that are of low magnitude versus high magnitude, we are more effective at discounting smaller losses (Stewart, Chater, & Brown, 2006). For example, compare the feeling of losing $2.00 to losing $200. We are more likely to be effective at discounting a loss of $2.00 compared to losing $200. On the other hand, when it comes to gains, people readily accept potential wins regardless of the magnitude.

For small amounts of money, small losses can be easily discounted. This results in the positive feelings associated with small positive outcomes outweighing the negative feelings associated with small negative outcomes. In other words, "gains loom larger than losses" (i.e., a reversal of loss aversion). In contrast, for large amounts of money, large negative outcomes are harder to discount and therefore cannot be outweighed by the positive feelings associated with large positive outcomes. Therefore, loss aversion is observed for large magnitude outcomes. Thus, this research demonstrates that differences in how we respond to gains and losses can depend on the magnitude of the outcome.

Learning asymmetry for wins and losses. Our previous work provided evidence for an asymmetry in gain and loss stimulus-outcome associations in a probabilistic value learning task (VLT; Lin et al. 2020). This evidence took two forms: 1) a meta-analysis 10 studies using a probabilistic learning task developed by Raymond & O'Brien (2009), or similar value learning task, and 2) results from two new empirical studies demonstrating that the learning asymmetry was evident whether wins and losses led to point or monetary earnings, and whether or not participants were informed about the outcome contingencies governing wins and losses. Of relevance is the fact that the magnitude of the win and loss payoffs in these studies was very modest.

In the studies adopting the VLT or similar probabilistic value learning tasks, that were included in our meta-analysis (Lin et al., 2020), the magnitude of wins and losses were of relatively small magnitude (e.g., $+/-0.50, +/- 5 points). Furthermore, the finding that the learning asymmetry is evident whether wins and losses lead to point or monetary earnings suggests that psychologically, the two types of outcomes are perceived similarly. The pattern of the learning asymmetry would coincide with the strategy of discounting in that participants may have easily discounted small losses and readily accepted small gains. Thus, participants prioritized learning the optimal choice for the win pair in order to maximize wins over learning the optimal choice for the loss pair to minimize losses. However, whether increasing the magnitude of the outcomes would either reverse or diminish this learning asymmetry remains unanswered. To our knowledge, the effect of outcome magnitude on learning win and loss associations has not been tested in the VLT. Therefore, the goal of the present study was to test the hypothesis that increasing the magnitude of win and loss outcomes may alter the learning pattern for win and loss associations.

To test this hypothesis, we varied the magnitude of win and loss outcomes and measured the effects on learning. Two win and two loss pairs were presented across five trial blocks in an intermixed and randomized order. One win pair and one loss pair led to a low magnitude outcome, with the same point value as our previous work (+/-5 points). The other win pair and loss pair led to a high magnitude outcome (+/- 50 points).

Our previous work has also demonstrated that explicit knowledge of valence and predictiveness of outcomes differ for win and loss associations (Lin et al. 2020). Therefore, we also included a memory task to examine how increasing the magnitude of wins and losses affects people's explicit memory for the outcomes associated with each image in the VLT.

Furthermore, we examined whether the difference in learning win and loss associations for low and high magnitude outcomes would be associated with trait sensitivity to punishment and reward and loss aversion. Participants completed the Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ; Torrubia, Avila, Molto, & Caseras, 2001) and a gambling task designed to measure the extent to which people weigh losses compared to equivalent gains (Kahneman & Tversky, 1979; Tom, Fox, Trepel, & Poldrack, 2007).

Lastly, as in previous chapters, we derived reinforcement learning parameters to characterize choice behavior in the VLT. The learning rate and inverse temperature parameter were compared between the low and high outcome magnitude condition to examine whether the learning rate and balance between exploration and exploitation differed between the two conditions.

Due to the COVID-19 pandemic, the present tasks were administered online, as described below. To our knowledge, this is the first time the VLT was run in an online experiment. Therefore, as a precursor to testing the effect of outcome magnitude on win and loss associations, the VLT used by Lin et al. (2020) was administered to participants recruited from an online platform to test whether the learning asymmetry replicates with an online sample.

Experiment 1

Method

Participants

Sixty adults (31 males, 27 females, 1 who did not disclose sex; with age range 18-40 years, *M* = 27.97, *SD* = 6.82) were recruited from Prolific. The minimum number of participants was informed by a power analysis (G*Power; Faul, Erdfelder, Buchner, & Lang, 2009). We over-recruited to account for counterbalancing considerations and participants who might fail to reach the learning criterion. All procedures were approved by the University of Michigan, Institutional Review Board. Participants were prescreened to exclude those who reported having mild cognitive impairment/dementia and who did not have normal or corrected-to-normal vision. Participants were compensated at a rate of $10 per hour. As with other studies administering the VLT, we adopted a learning criterion in order for participants to be included in the data analysis. Participants should select the optimal choice on at least 65% of the trials in the final block of the VLT for both the win and loss pair. The final sample used for data analysis was 37 participants (17 males, 20 females, with age range 18-39 years, *M* = 27.38 years, *SD* = 7.21).

Materials and Procedure

Participants completed the VLT and Memory Task in Pavlovia, which were programmed using PsychoPy v3.0.

Stimuli. Eighteen grayscale images of landscape scenes were used as the stimuli for the VLT and Memory Task. Six of these images were used for the VLT and Memory Task, and 12 were used as new images for the Memory Task (Rissman, Gazzaley, & D'Esposito, 2009).

Value Learning Task. The VLT was designed to replicate the version presented by Lin et al. (2020). Therefore, the parameters of the task are the same. However, for reference, a description of the task is also presented here. The task consisted of six scenes, which were divided into three pairs: one win pair, one loss pair, and one no-change pair. The pairing of scenes, assignment of each scene to a valence (win versus low) and probability (20% versus

80%), was counterbalanced across participants. Each image pair was presented 100 times in a randomized manner, across five blocks, yielding a total of 300 trials across all pairs.

On each trial, a scene pair appeared on the screen, one above and one below a central fixation cross (Figure 4.1). Each scene was always presented with its pair, but the location of the scene (top vs. bottom) was randomized from trial to trial. For each trial, participants selected one scene from the pair, with the goal of maximizing the number of points they earned. Participants pressed the "f" key to select the top scene or "j" key to select the bottom scene. For the win pair, selection of a scene resulted in either a win or no change; for the loss pair, selection of a scene resulted in either a loss or no change, depending on the contingencies governing the outcome; for the no change pair, selecting of either scene always resulted in a no change in points.

Figure 4.1. Schematic of the Value Learning Task.
On each trial, a pair of neutral landscape scenes is presented, one above a central fixation cross, and the other below. The scene pair remained on the screen until participants made a selection. Once participants made a selection, a yellow border

appeared around the screen, which remained for 500 msec. Feedback was provided for each selection, which remained on the screen for 1500 msec. ITI was 500 msec.

The probability of a win, loss, or no change outcome varied for the scenes constituting the win and loss pair. One scene produced the corresponding win or loss outcome with a high probability (80%) while its counterpart produced its assigned win or loss outcome with a low probability (20%). The assigned probability governing each outcome was constant for the entire 5-block session. To maximize points, participants had to learn the 'optimal' scene for each pair. For the win pair, the optimal scene is the scene which has a higher probability of gaining points (80% win scene). Whereas the suboptimal scene has a lower likelihood of gaining points (20% win scene). For the loss pair, the optimal scene is the scene which has a lower likelihood of losing points (20% loss scene). Whereas the suboptimal scene has a higher likelihood of losing points (80% loss scene).

When participants chose a scene, feedback was displayed with a message "Win" in green, "Lose" in red, or "No change" in black, depending on the scene pair just presented and the probability governing the outcome. A running total of the earnings appeared in the top right-hand corner of the screen. The VLT was designed to take participants approximately 25-30 minutes to complete.

Memory Task. This task was designed to test participant's explicit knowledge of the outcome associated with each scene from the VLT. The Memory Task consisted of the 6 scenes shown in the VLT along with 12 new scenes. Each scene was presented one at a time at the center of the computer screen. The scenes from the VLT were each presented four times, whereas new scenes were each presented twice in a randomized order yielding a 50% probability of the old versus new items. For each scene, participants indicated which outcome the scene was most likely to be associated (very likely to lead to a win, occasional win, no change, occasional loss, or very likely to lead to a loss), or whether it was a new image (Figure 4.2). Completion of the memory task was self-paced, and no feedback was given about performance on this task. The Memory Task was designed to take approximately 5 minutes to complete.

1 = Very likely to win, 2= Occasionally win,
3 = No change,
4 = Occasionally lose, 5 = Very likely to lose,
6 = None (new image)

Figure 4.2. Schematic of the Memory Task.
This task consists of the 6 scenes from the Value Learning Task and 12 new scenes. Participants select the outcome that was most likely associated with each scene.

Results and Discussion

Value Learning Task

The probability of optimal choice was compared between the win and loss pairs. The no change pair was examined separately and discussed below. Probability of optimal choice was calculated as the proportion of trials for which participants chose the high-probability win scene (80% win) for the win pair and low-probability loss scene (20% loss) for the loss pair. The probability of optimal choice was calculated for each set of 20 trials, resulting in five 20-trial blocks. Mean probability of optimal choice for the win pair and loss pair were then compared using a 5 (Trial Block: 1-5) x 2 (Outcome Valence: win, loss) within-subjects ANOVA. Greenhouse-Geisser corrections were applied when assumptions of sphericity were violated, and

p-value significance cutoffs were adjusted using a Bonferroni correction (α /number of comparisons) to account for multiple pairwise comparisons.

A significant main effect emerged for Valence, revealing that participants' optimal choice was higher for the win pair when compared with the loss pair, $F(1, 36) = 7.06$, $p = .01$, $\eta_p^2 = .16$ (Figure 4.3). A significant main effect also emerged for Trial Block, $F(3.07, 110.63) = 29.64$, $p < .001$, $\eta_p^2 = .45$. Follow-up pairwise comparisons showed that probability of optimal choice increased from Trial Block 1 to Trial Block 2, $p < .001$, but did not significantly differ in subsequent blocks. No significant interaction emerged between Outcome Valence and Trial Block, $F(2.30, 82.76) = 0.92$, $p = .46$.[7]

For the no-change pair, one no-change scene was arbitrarily selected as the correct scene. The probability of selecting the scene was then examined by a one-way repeated-measures ANOVA as a function of Trial Block (1-5). The probability of selecting the chosen "correct" no-change scene was .48 (*SD* = .30) across all participants and did not significantly differ by Trial Block, $F(3.43, 202.22) = .27$, $p = .87$. Thus, there were no systematic biases in selecting a no-change scene.

---

[7] When examining performance on the VLT for participants who did not reach the learning criterion, the results showed no main effect for Outcome Valence, $p = .20$, or Trial Block, $p = .96$, and no significant interaction between Outcome Valence and Trial Block, $p = .34$. While these results suggest that p(optimal choice) did not differ between the win and loss pairs, performance was at chance level for the win pair ($M = .49$, $SE = .03$) and loss pair ($M = .55$, $SE = .03$). Thus, it appears that participants did not learn the optimal choice, limiting the conclusion that learning was similar between win and loss pairs.

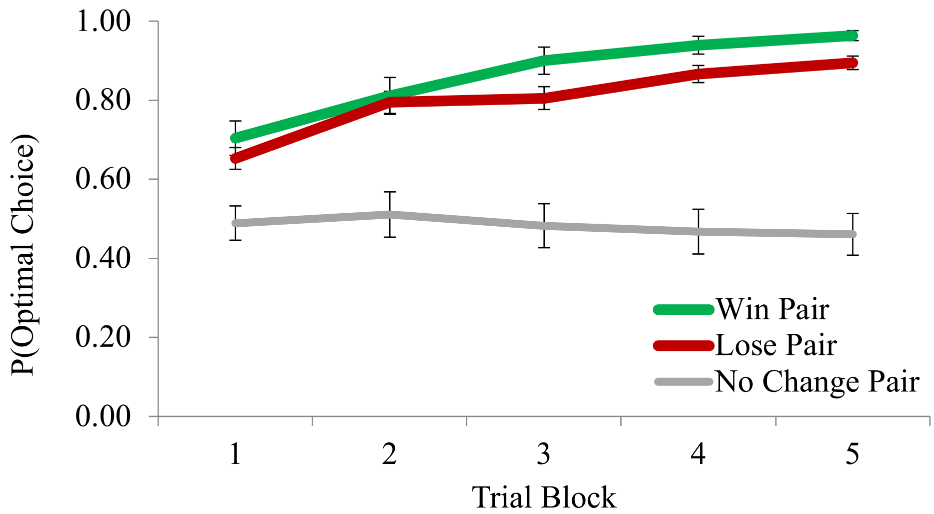


Figure 4.3. Performance on the Value Learning Task for participants who reached the learning criterion.
Participants selected the optimal choice for the win pair at a higher rate than the loss pair.

The results from the VLT replicate the learning asymmetry reported by Lin et al. (2020). This replication provides an important foundation for subsequent studies administering the VLT to an online sample of participants. More specifically, if any changes in the learning pattern are observed with manipulations to the task parameters, then an alternative explanation that the observed changes are attributed to the difference in an online sample versus a sample of participants completing the task in a research laboratory can be ruled out.

Memory Task

To examine performance on the Memory Task, we first used a binary logistic regression to examine accuracy in outcome attribution. Valence (win versus loss, with win coded as 0) and Optimality (suboptimal versus optimal, with suboptimal coded as 0) were entered as predictors of outcome accuracy (incorrect outcome coded as 0, correct outcome coded as 1). The interaction term, Valence x Optimality significantly increased model fit. Therefore, the final model included Valence, Optimality, and Valence x Optimality as predictors of outcome accuracy.

The results showed that Valence was a significant predictor of memory accuracy, indicating that memory accuracy was higher for loss scenes when compared with win scenes (Table 4.1). However, the Valence x Optimality interaction suggests that this pattern differed between optimal and suboptimal scenes. For optimal scenes, memory accuracy was higher for win scenes (80% win) when compared with loss scenes (20% loss) (Figure 4.4). On the other hand, for suboptimal scenes, memory accuracy was higher for loss scenes (80% loss) when compared with win scenes (20% win). Optimality was also a significant predictor of memory accuracy, indicating that memory accuracy was higher for optimal scenes when compared with suboptimal scenes. The Valence x Optimality interaction revealed that this pattern held for both win and loss scenes but was more pronounced for win scenes.

Table 4.1.
Binary Logistic Regression Predicting Memory Accuracy from Valence and Optimality.

Predictor	β	OR	95% CI for OR
Intercept	-2.00***	0.14	(3.71, 12.15)
Valence	2.28***	9.77	(5.05, 20.14)
Optimality	4.95***	141.18	(59.84, 369.13)
Valence*Optimality	-4.16***	0.02	(0.01, 0.04)

Note: OR = Odds Ratio.
***$p < .001$, **$p < .01$, *$p < .05$

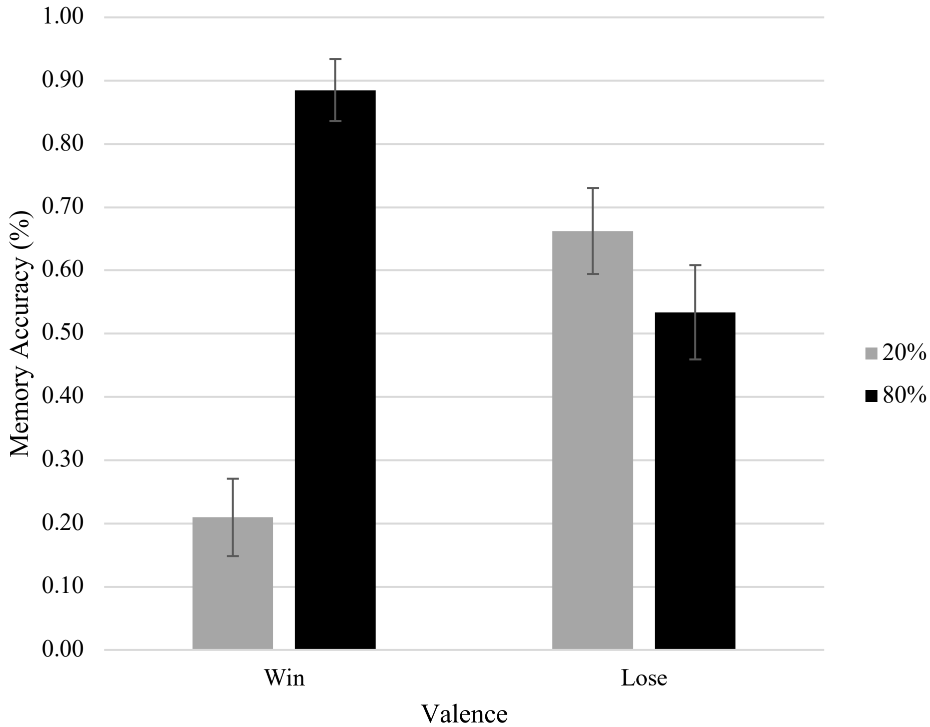


Figure 4.4. Accuracy of explicit knowledge of outcomes associated with the VLT scenes

Outcome Attributions. We also examined the range of participants' responses on the Memory Task for each VLT scene using a multinomial logistic regression. Valence, Probability, and the interaction between Valence and Probability were the predictor variables, and the probability that each response option was selected was the outcome variable (Table 4.2).

For the 80% win scene, "very likely to win" was selected most frequently as the outcome associated with the scene (Figure 4.5). Misattributions were rare. For the 20% win scene, participants were more likely to select "no change," followed by "occasionally win" and "occasionally lose." Thus, for the 20% win scene, participants made valence misattributions but were unlikely to attribute either a high probability win or loss.

For the 80% loss scene, participants were more likely to select “very likely to lose,” followed by “occasionally lose” (Figure 4.6). For the 20% loss scene, participants were more likely to select “occasionally lose.” Other attributions were primarily “no change” or “occasionally win.” Thus, for the 80% loss scene, participants confused the probability of a loss but were unlikely to misattribute the scene to a win. For the 20% loss scene, misattributions included a low probability win. However, participants were unlikely to attribute the scene to a high probability win or loss.

Table 4.2.
Multinomial Logistic Regression Predicting Outcome Response for each VLT scene from Valence and Probability for Learners.

Outcome Response	Valence	Probability	Valence * Probability
Very likely to win	0.00	1120.38***	118.72
Occasionally win	2.66*	6.03*	0.01***
Occasionally lose	17.75***	1.02	0.45
Very likely to lose	3.23	0.00	1.21×10^5
None (new)	2.59	9.15**	0.05**

Note. Values in the table are odds ratio. “No change” was selected as the reference level. Valence was coded as 0 for the win scenes, and 1 for the loss scenes. Probability was coded as 0 for low probability, and 1 for high probability.
p* < .05, ** *p* < .01, * *p* < .001

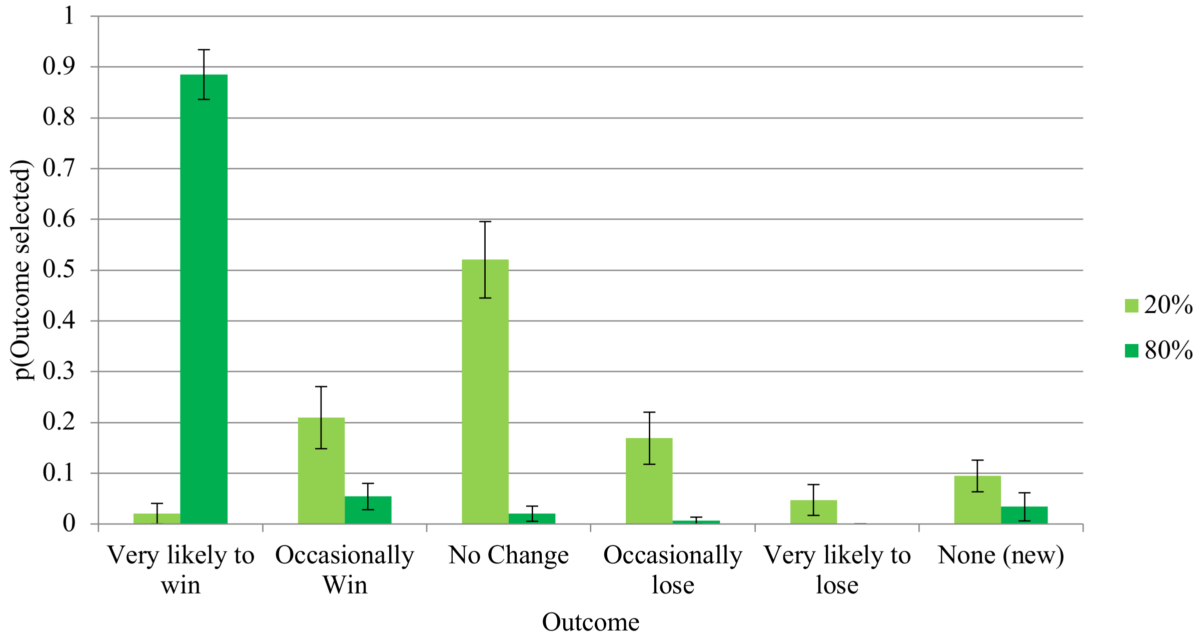


Figure 4.5. Outcome attributions for the win scenes.

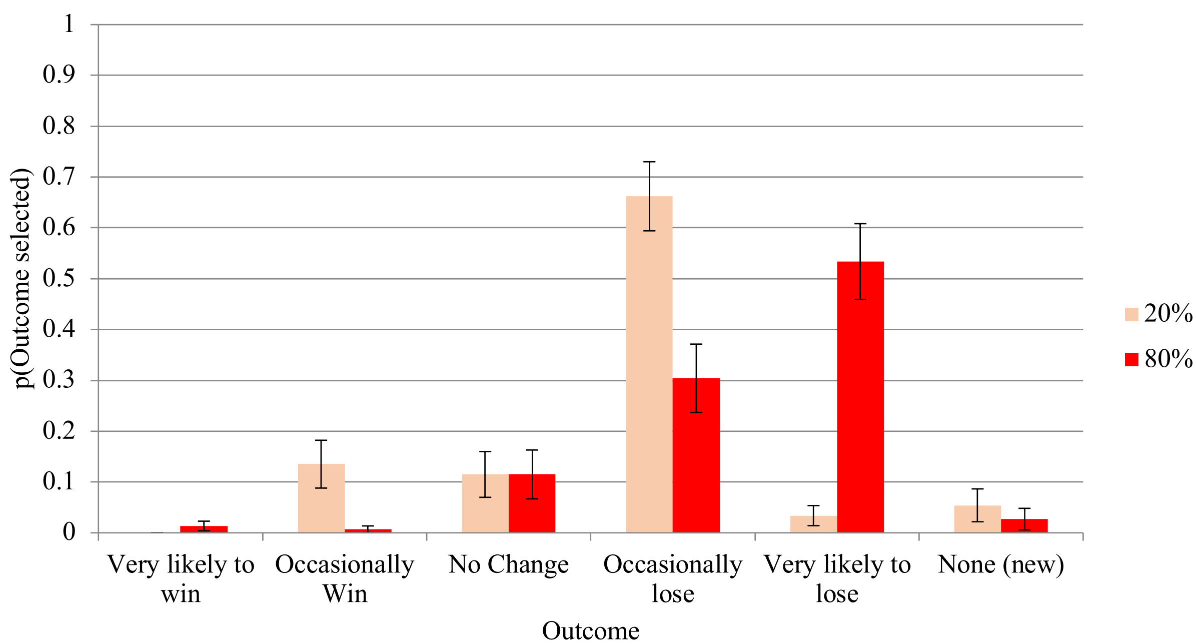


Figure 4.6. Outcome attributions for the loss scenes.

The results from the memory task demonstrate that participants had explicit knowledge of the optimal choice for each win and loss pair. For optimal scenes, participants showed better explicit knowledge of win scenes compared to loss scenes. Whereas for suboptimal scenes, participants showed better explicit knowledge of loss scenes compared to win scenes. This pattern of results replicates the findings from Lin et al. (2020). Therefore, in the next experiment, the effect of outcome magnitude was evaluated on learning and explicit knowledge of the outcomes associated with each scene. We added an additional win and loss pair to the VLT in order to have one win and loss pair in the low outcome magnitude condition and a second win and loss pair to the high outcome magnitude condition. We excluded the "no change" condition to offset the increase in number of trials that resulted from adding two new pairs.

Experiment 2

Method

Participants

One-hundred-sixty-eight adults (85 males, 80 females, 3 who did not disclose sex; with age range 18-41 years, $M = 27.86$, $SD = 6.35$) were recruited from Prolific. The minimum number of participants was informed by a power analysis (G*Power; Faul, Erdfelder, Buchner, & Lang, 2009) with over-recruiting to account for counterbalancing and to anticipate participants who might fail to reach the learning criterion. All procedures were approved by the University of Michigan, Institutional Review Board. Participants were prescreened to exclude those who reported having mild cognitive impairment/dementia and who did not have normal or corrected-to-normal vision. Participants were compensated at a rate of $10 per hour plus a bonus of up to $15 based on their performance on the VLT and Loss Aversion Task described below. We also adopted a learning criterion for the win and loss conditions. To be included in our analyses as "learners", participants needed to select the optimal choice on at least 65% of the trials in the final block of the VLT for both win pairs and loss pairs. This resulted in excluding 97 participants (non-learners: 45 males, 51 females, 1 who did not disclose sex, with age range 18-41 years, M = 28.13 years, SD = 6.44).[8] The final sample used for data analysis was 71 participants (learners: 40 males, 29 females, 2 who did not disclose sex, with age range 18-40 years, M = 27.51 years, SD = 6.25). For completeness, the analyses presented below were repeated for non-learners, which can be found in the Supplemental Material.

Materials and Procedure

After providing consent, participants completed a questionnaire using Qualtrics. They were then redirected to Pavlovia to complete the VLT, Memory Task, and Loss Aversion Task, described below, which were programmed using PsychoPy v3.0.

Questionnaire. Participants completed the Sensitivity to Punishment and Sensitivity to Reward questionnaire (SPSRQ), which was designed to measure how individuals respond to reward and punishment (Torrubia, Avila, Molto, & Caseras, 2001). This 48-item questionnaire consists of statements describing behaviors that an individual may engage in. Participants indicated whether the behaviors are descriptive of them by selecting either "*yes*" or "*no*" for each statement. Scores are created by adding all the "yes" answers to items corresponding to the

[8] A series of chi-square tests of independence were conducted to examine whether learning a pair of a given valence was associated with learning of another pair. A test of independence was conducted between win pairs versus loss pairs for the low-magnitude condition, win pairs versus loss pairs for the high-magnitude condition, low-magnitude win pair versus high-magnitude win pair, and low-magnitude loss pair versus high-magnitude loss pair. Across all tests, participants who were categorized as a non-learner for a given pair were also categorized as a non-learner for other pairs, ps < .001. Thus, there was no evidence that participants tended to be excluded for not learning a given valence (e.g., win or loss) or magnitude condition (e.g., low versus high) versus its counterpart (e.g., loss or win; high versus low).

Sensitivity to Punishment and Sensitivity to Reward subscales. The completion time for the SPSRQ was approximately 5 – 10 minutes.

Stimuli. Twenty-four grayscale images of landscape scenes were used as the stimuli for the VLT and Memory Task. Eight of these images were used for the VLT and Memory Task, and 16 were used as new images for the Memory Task (Rissman, Gazzaley, & D'Esposito, 2009).

Value Learning Task. The VLT reported in Experiment 1 was modified in the following ways: 1) an additional win and loss pair were added, 2) the no change pair was removed to offset the increase in trials that resulted from adding two new pairs, 3) a colored border was introduced around the perimeter of the computer screen throughout all phases of a trial to signal the outcome magnitude, and 4) a performance-based incentive was introduced to make the difference in outcome magnitude more salient. These modifications are described in more detail below.

The VLT consisted of eight scenes, which were divided into four pairs: two-win pairs, two-loss pairs. One pair within each valence was randomly assigned to a low outcome magnitude condition (+/- 5 points), while the other pair was assigned to a high outcome magnitude condition (+/- 50 points). The pairing of scenes, assignment of each scene to a valence (win versus low) and probability (20% versus 80%), and assignment of each pair to either the low or high outcome magnitude condition was counterbalanced across participants. Each image pair was presented 100 times in a randomized manner, across five blocks, yielding a total of 400 trials across all pairs.

When participants chose a scene, feedback was displayed with a message “Win” in green, “Lose” in red, or “No change” in black, depending on the scene pair just presented and the probability governing the outcome (Figure 4.7). The point value for wins and losses (+/- 5 or +/- 50) also appeared on the center of the screen. A running total of the earnings appeared in the top right-hand corner of the screen.

Prior to beginning the VLT, participants were informed that the stakes for some trials may be higher than others, which would be highlighted using a colored border (purple or turquoise). The colored border remained on the screen throughout all phases of a trial (scene presentation, scene selection, and feedback). The assignment of the colored border to low versus high stakes trials was counterbalanced across participants. Furthermore, participants were informed that their performance on the task would lead to a monetary bonus. The performance-based monetary incentive was introduced to make the effect of low versus high magnitude outcomes more relevant to participants. The VLT required participants approximately 30-40 minutes to complete.

Figure 4.7. Schematic of the Value Learning Task in the Outcome Magnitude Experiment.

In this trial, a purple border was indicative of a low-stakes trial (+5 points). On each trial, a pair of neutral landscape scenes is presented, one above a central fixation cross, and the other below. The scene pair remained on the screen until participants made a selection. Once participants made a selection, a yellow border appeared around the screen, which remained for 500 msec. Feedback was provided for each selection, which remained on the screen for 1500 msec. ITI was 500 msec.

Memory Task. The memory task reported in Experiment 1 was modified to account for the additional scenes in the VLT. The task consisted of 8 VLT scenes (each image repeated 4 times) and 16 new scenes (each image repeated 2 times). Furthermore, we examined participants' explicit knowledge of the outcome magnitude. If a participant indicated that the scene appeared in the VLT (i.e., any response except for 'none/new'), then participants were asked whether the scene was associated with a low-stake or high-stake outcome (Figure 4.8). Completion of the memory task was self-paced, and no feedback was given about performance on this task. The Memory Task required participants 5-10 minutes to complete.

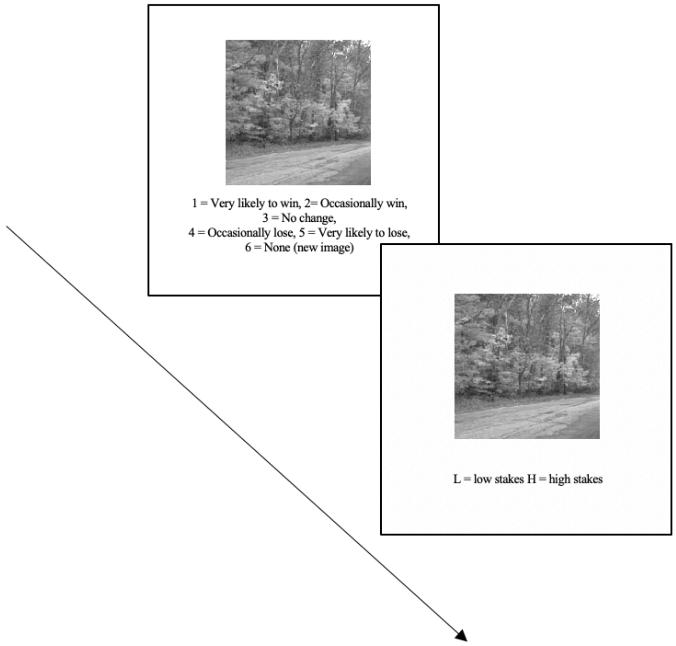


Figure 4.8. Schematic of the Memory Task in the Outcome Magnitude Experiment. Participants indicated with what outcome the scene was most likely associated. Any response other than "none" was followed by a second screen asking participants to indicate whether the scene was associated with a low-stake or high-stake outcome.

Loss Aversion Task. After participants finished the Memory task, they completed the Loss Aversion Task, which was designed to measure the extent to which people weigh losses more than objectively equivalent gains (i.e., loss aversion; Kahneman & Tversky, 1979; Tom, Fox, Trepel & Poldrack, 2007). In this gambling task, participants are presented with a randomized series of hypothetical gambles in which they have a 50% chance of winning or 50% chance of losing an amount presented on the screen (e.g., 50% of winning $24, 50% of losing $14). The potential gains ranged from $10 to $30, while potential losses ranged from $5 to $15. For each gamble, participants decided whether to accept or reject the proposed gamble.

Prior to beginning the task, participants were informed that at the end of the task one of the gambles would be randomly selected. If the randomly selected gamble was one that they accepted, the participant would win or lose the amount displayed based on the outcome of a coin flip. The outcome of the final gamble added or subtracted from the bonus they earned from the VLT. The Loss Aversion Task took 5 – 10 minutes to complete.

Results

Value Learning Task

To examine whether Outcome Magnitude had an effect on the presence or absence of a learning asymmetry for wins and losses, a 5 (Trial Block: 1-5) x 2 (Outcome Valence: win versus loss) x 2 (Outcome Magnitude: low versus high) within-subjects ANOVA was conducted with probability of optimal choice as the dependent variable. Greenhouse-Geisser corrections were applied when assumptions of sphericity were violated. The *p*-value significance cutoffs for follow-up pairwise comparisons were adjusted using a Bonferroni correction (α/number of comparisons).

The results showed a significant main effect of Outcome Valence indicating that percent of optimal choice was higher for the win pairs (M = .87, SE = .01) versus loss pairs (M = .81, SE = .01), $F(1, 70) = 20.89$, $p < .001$, n_{p}^2 = .23. A significant main effect of Trial Block emerged indicating that percent of optimal choice increased with each subsequent block, $F(2.97, 207.71) = 172.42$, $p < .001$, n_{p}^2 = .71. A significant main effect of Outcome Magnitude emerged indicating that percent of optimal choice was higher for scene pairs that led to outcomes of high magnitude (M = .87, SE = .01) versus low magnitude outcomes (M = .82, SE = .01), $F(1, 70) = 22.18$, $p < .001$, n_{p}^2 = .24.

When evaluating the two-way interactions, a significant interaction emerged between Outcome Magnitude x Trial Block, $F(3.02, 211.09) = 11.11$, $p < .001$, n_{p}^2 = .14. Follow-up pairwise comparisons indicated that optimal choice was higher for the high magnitude pairs versus low magnitude pairs during Trial Blocks 1-3, *p*s $< .003$, but did not significantly differ between Trial Blocks 4-5, *p*s $> .28$ (Figure 4.9). Thus, by the end of the task, participants showed similar learning for pairs that led to low and high magnitude outcomes. No significant interactions emerged between Valence x Outcome Magnitude, $F(1, 70) = 2.74$, $p = .10$. However, since evaluating this interaction directly tests our hypothesis, we conducted a paired samples *t*-test to compare probability of optimal choice between the win and loss pair for each magnitude condition. The results showed that learning was better for the win versus loss pair for both the low magnitude condition, $t(70) = 2.05$, $p = .04$, and high magnitude condition, $t(70) = 5.37$, $p < .001$ (Figure 4.10). No significant interaction emerged between Valence x Trial Block $F(2.75, 192.59) = 1.22$, $p = .30$. Lastly, the three-way interaction between Valence x Outcome Magnitude x Trial Block was not significant, $F(2.71, 189.39) = 0.66$, $p = .56$. However, to further test our hypothesis, we conducted a series of exploratory analyses to compare learning for wins and losses throughout all trial blocks in each outcome magnitude condition. The results showed that for the low magnitude condition, probability of optimal choice was higher for the win pair versus loss pair only during Trial Block 4 and did not significantly differ for any other trial blocks (Figure 4.11). For the high magnitude condition, probability of optimal choice was higher for the win pair versus loss pair throughout all Trial Blocks, *p*s $< .002$. Thus, the learning asymmetry favoring wins emerged for all trial blocks for the scenes that led to high magnitude outcomes. On the other hand, for the low magnitude condition, participants displayed similar learning between the win and loss pair throughout most of the trial blocks, reaching a similar level of learning by the end of the task.

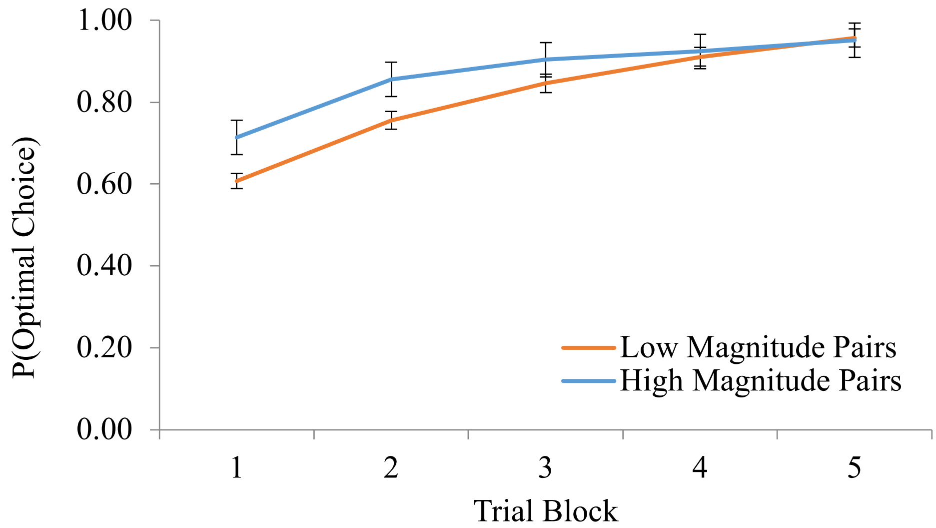


Figure 4.9. Interaction between Outcome Magnitude and Trial Block. Optimal choice was significantly higher for the high magnitude pairs versus low magnitude pairs during Trial Blocks 1-3 but did not significantly differ during Trial Blocks 4-5.

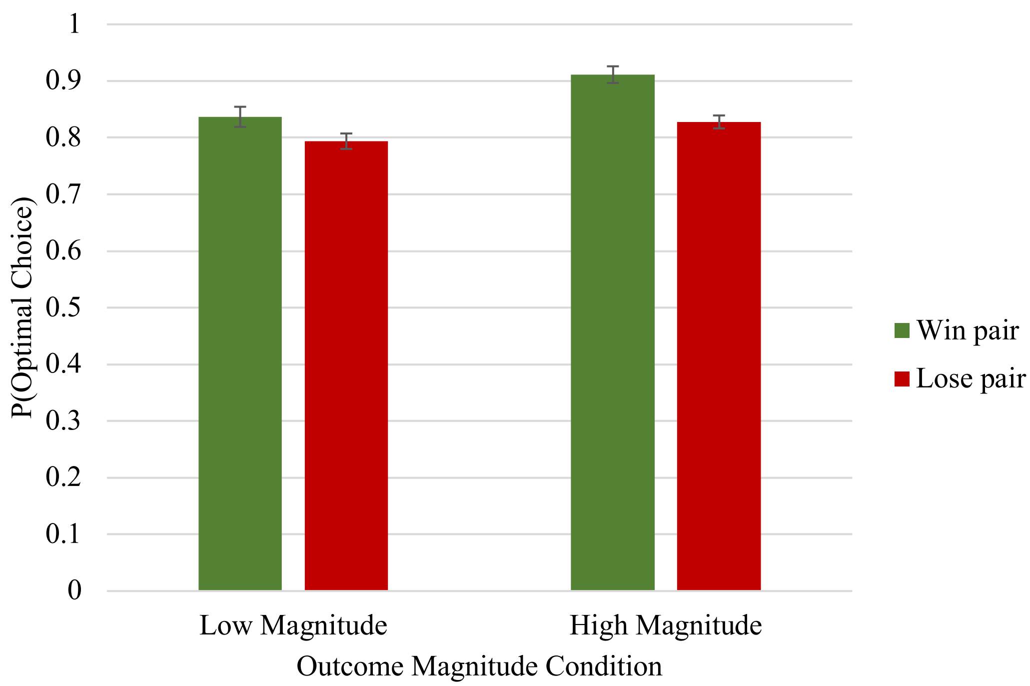


Figure 4.10. Probability of optimal choice was significantly higher for the win pair versus loss pair in each outcome magnitude condition.

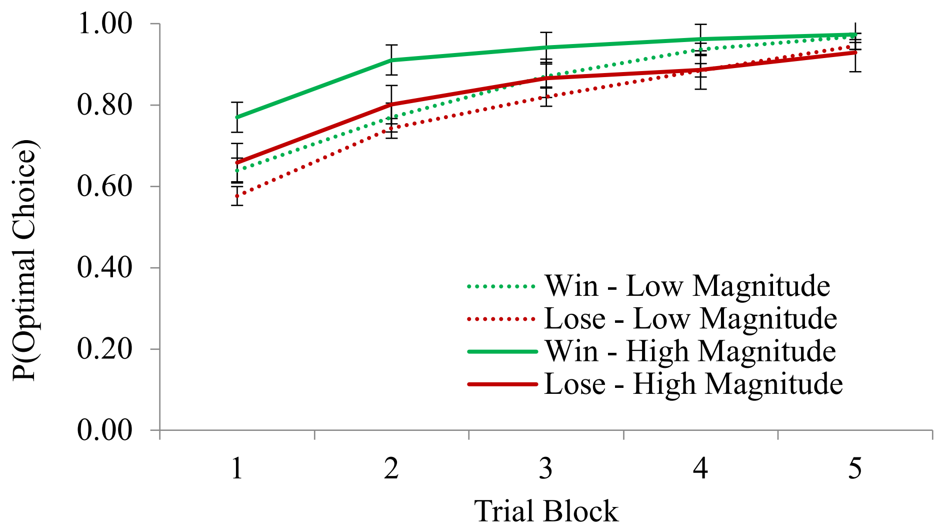


Figure 4.11. Performance on the VLT for participants who met the learning criterion. Probability of optimal choice is displayed for the win and loss pair in each outcome magnitude condition.

Correlations between Questionnaire Measures and Performance on the VLT

To examine whether measures of sensitivity to punishment and reward and loss aversion were associated with the learning difference between win and loss outcome associations, a series of correlations were conducted. A learning difference score was calculated for each magnitude condition by subtracting percent of optimal choice across all five trial blocks for the loss pair from the win pair for each participant. The sign of the difference score was preserved. Therefore, positive values indicate wins were learned better than losses. Whereas negative values indicate losses were learned better than wins. The *p*-value cutoff for significance was adjusted to account for the number of correlations using the Holm-Bonferroni method. The results showed that trait sensitivity to punishment and reward and loss aversion were not significantly correlated with the learning difference for win and loss pairs for either the low magnitude or high magnitude condition, *p*s > .05 (Table 4.3).

Table 4.3.
Pearson r values for the Correlation between the Learning Difference Score and Measures of Sensitivity to Punishment and Reward and Loss Aversion.

Learning Difference	Sensitivity to Punishment	Sensitivity to Reward	Loss Aversion
Low magnitude pairs	.12	.03	.02
High magnitude pairs	.14	-.13	.07

Note. The learning difference score was calculated for each magnitude condition as p(optimal choice) averaged across all trials for the win pair minus loss pair.
$^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$

Memory Task

Memory accuracy for VLT Scenes. We used a binary logistic regression analysis to examine explicit knowledge of the outcomes associated with each VLT scene. Valence (win versus loss, with win coded as 0), Optimality (suboptimal versus optimal, with suboptimal coded as 0), and Outcome Magnitude (low versus high, with low coded as 0) were entered as predictors of outcome accuracy (incorrect outcome coded as 0, correct outcome coded as 1). Interaction terms were added iteratively to examine whether each term increased model fit. The final model only included interaction terms which significantly increased model fit. Therefore, the final model included Valence, Optimality, Magnitude, and Valence x Optimality as predictors of outcome accuracy (Table 4.4).

The results showed that Valence was a significant predictor of outcome attribution accuracy. Accuracy was higher for loss scenes when compared with win scenes. Optimality was also a significant predictor of outcome attribution accuracy. Accuracy was higher for optimal scenes when compared with suboptimal scenes. Outcome Magnitude was also a significant predictor of outcome attribution accuracy. Overall, accuracy was higher for scenes that led to high magnitude outcomes (M = .54, SD = .20) versus low magnitude outcomes (M = .33, SD = .21). The interaction between Valence and Optimality was also significant. The interaction revealed that for win scenes, accuracy was higher for optimal scenes when compared with suboptimal scenes (Table 4.5; Figure 4.12). Whereas for loss scenes, accuracy did not significantly differ between optimal and suboptimal scenes. Furthermore, for suboptimal scenes, accuracy was higher for loss scenes when compared with win scenes. Whereas for optimal scenes, accuracy was higher for win scenes when compared with loss scenes.

Table 4.4.
Binary Logistic Regression Predicting Memory Accuracy from Valence, Optimality, and Outcome Magnitude.

Predictor	β	OR	95% CI for OR
Intercept	-2.98***	0.05	(0.03, 0.07)
Valence	2.07***	7.89	(5.67, 11.13)
Optimality	3.65***	38.47	(26.82, 56.06)
Outcome Magnitude	1.27***	3.55	(2.87, 4.42)
Valence*Optimality	-3.78***	0.02	(0.01, 0.04)

Note: OR = Odds Ratio.
****p* < .001, ***p* < .01, **p* < .05

Table 4.5.
Means and Standard Deviation of Memory Accuracy for the VLT scenes.

	Low Magnitude		High Magnitude	
	20%	80%	20%	80%
Win	0.11(0.26)	0.60(0.45)	0.15(0.31)	0.89(0.26)
Lose	0.34(0.42)	0.26(0.39)	0.49(0.45)	0.63(0.42)

Note. For the win pairs, the optimal scene is the 80% win scene, the suboptimal scene is the 20% win scene. For the loss pairs, the optimal scene is the 20% loss scene, the suboptimal scene is the 80% loss scene.

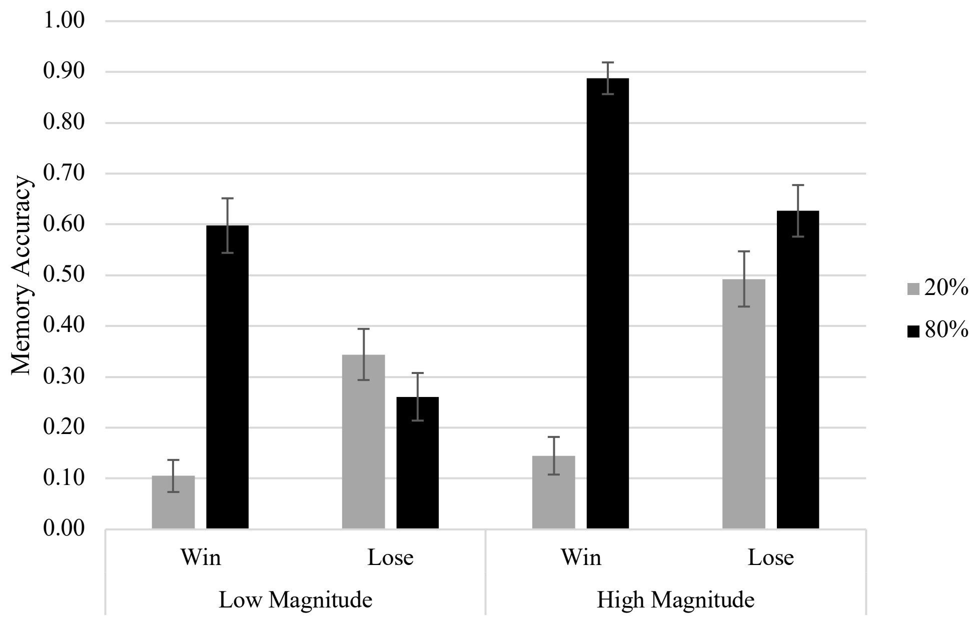


Figure 4.12. Accuracy of explicit knowledge of outcomes associated with the VLT scenes.

Outcome Attributions. We also examined the range of participants' responses on the Memory Task for each VLT scene using a multinomial logistic regression, with all possible response options for the memory task as the outcome variable, and Valence, Probability, and Outcome Magnitude as predictor variables (Table 4.6).

Outcome attributions for win pairs. For both low and high magnitude conditions, participants were more likely to correctly attribute "Very likely to win" as the outcome associated with the 80% win scene compared to other responses (

Figure 4.13) For the high magnitude pair, misattribution errors were rare for the 80% win scene, whereas for the low magnitude pair, participants confused the probability of the win. For the 20% win scene in both the low and high magnitude condition, participants were more likely to attribute "No Change" as the outcome most likely associated with the scene. Misattribution errors included confusing the probability of the win and valence misattributions.

Outcome attributions for loss pairs. When examining responses for the 80% loss scene in the high magnitude condition, participants were more likely to correctly attribute “Very likely to lose” as the outcome most likely associated with the scene. Misattribution errors were mainly confusion in the probability of losses. Whereas for the 20% loss scene in the low magnitude, participants were more likely to select “occasionally lose” as the outcome associated with the scene. Misattribution errors were mainly confusion in the probability of losses.

For the 20% loss scene in the high magnitude condition, participants were more likely to attribute “occasionally lose” as the outcome most likely associated with the scene. On the other hand, for the 20% loss scene in the low magnitude condition, participants were just as likely to attribute “occasionally lose” and “no change” as the outcome associated with the scene. Misattribution errors included valence errors, although participants were unlikely to indicate the 20% loss scene had a high likelihood of resulting in a win.

In summary, these results revealed two general patterns: 1) memory accuracy was higher for scenes that led to high magnitude outcomes, 2) memory accuracy was higher for scenes that had a high probability of resulting in either a win or loss.

Table 4.6.
Multinomial Logistic Regression Predicting Outcome Response for each VLT scene from Valence and Probability.

Outcome Response	Valence	Probability	Outcome Magnitude	Valence * Probability	Valence * Magnitude	Probability * Magnitude	Valence * Probability * Magnitude
Very likely to win	.42	1677.15***	1.28	0.00	5.40	1.16	1.64×10^4
Occasionally win	2.52***	176.67***	1.77*	0.00***	0.51	0.13	15.61*
Occasionally lose	2.03**	7.23*	0.87	0.41	2.84**	0.16	2.24
Very likely to lose	0.15***	3.70	1.44	1.15*	6.02**	0.35	1.36
None (new)	0.93	36.14***	2.76***	0.11*	0.82	0.16	0.68

Note. Values in the table are odds ratio. “No change” was selected as the reference level. Valence was coded as 0 for the win scenes, and 1 for the loss scenes. Probability was coded as 0 for low probability, and 1 for high probability. Outcome magnitude was coded as 0 for low magnitude, and 1 for high magnitude.
p* < .05, ** *p* < .01, * *p* < .001

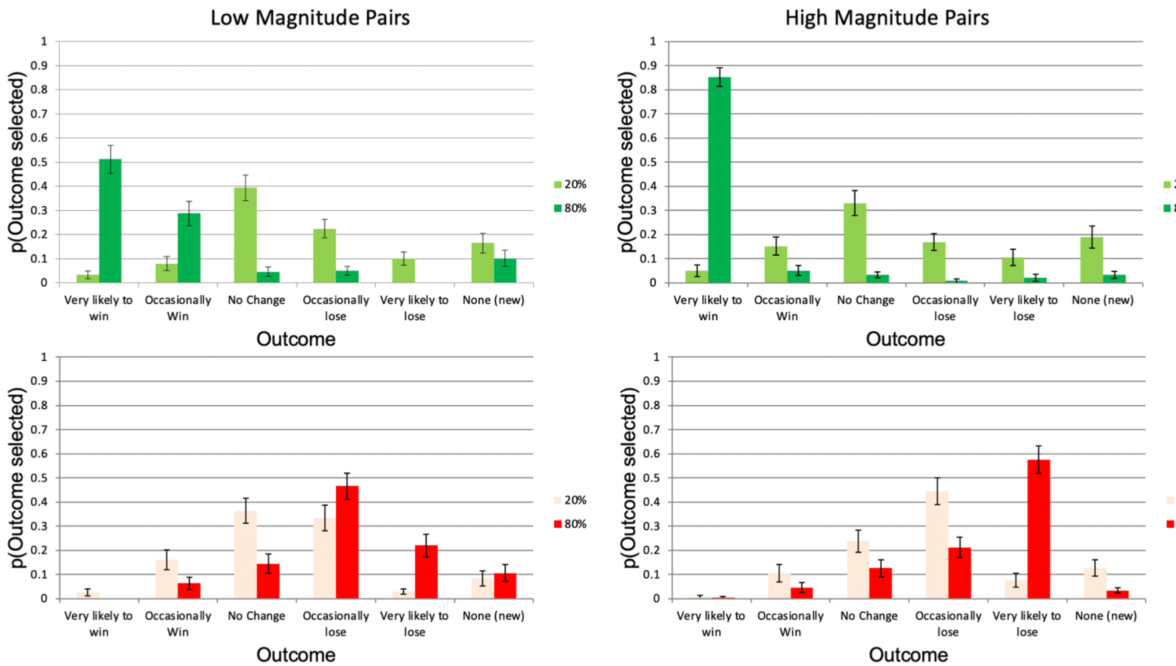


Figure 4.13. Outcome attributions in the Memory Task.
For each VLT scene, participants were instructed to indicate the outcome most likely associated with the scene.

Explicit knowledge of outcome magnitude. We also examined accuracy in attributing the outcome magnitude of scenes for which participants correctly indicated appeared in the VLT. We used a binary logistic regression analysis with Valence (win versus loss, with win coded as 0), Optimality (suboptimal versus optimal, with suboptimal coded as 0), and Outcome Magnitude (low versus high, with low coded as 0) as predictors of outcome magnitude accuracy (incorrect magnitude coded as 0, correct magnitude coded as 1). Interaction terms were added iteratively to examine whether each term increased model fit with the final model only including interaction terms which significantly increased model fit. Therefore, the final model included

Valence, Optimality, Magnitude, and Valence x Optimality as predictors of outcome magnitude accuracy (Table 4.7).

The results revealed that Valence was a significant predictor of outcome magnitude accuracy. Participants were more accurate at identifying the magnitude of outcomes associated with loss scenes versus win scenes. However, the Valence x Optimality interaction suggests that this pattern differed between optimal and suboptimal scenes (Figure 4.14). For suboptimal scenes, accuracy was higher for loss scenes versus win scenes. For optimal scenes, accuracy was higher for win scenes versus loss scenes. Optimality was also a significant predictor, suggesting that participants were more accurate at identifying the magnitude of outcomes associated with optimal scenes when compared with suboptimal scenes. However, the Valence x Optimality interaction suggests that this pattern differed between win and loss scenes. For win scenes, accuracy was higher for optimal scenes versus suboptimal scenes. Whereas for loss scenes, accuracy did not significantly differ between optimal and suboptimal scenes. Lastly, Outcome Magnitude was also a significant predictor of accuracy. Overall, participants were more accurate at identifying the magnitude of outcomes for low magnitude scenes when compared with high magnitude scenes.

Table 4.7.
Binary Logistic Regression Predicting Accuracy of Outcome Magnitude from Valence, Optimality, and Outcome Magnitude.

Predictor	β	OR	95% CI for OR
Intercept	2.61***	13.58	(7.95, 24.52)
Valence	0.40*	1.49	(1.00, 2.22)
Optimality	1.77***	5.85	(3.56, 9.90)
Outcome Magnitude	-1.00***	0.37	(0.27, 0.50)
Valence*Optimality	-1.92***	0.15	(0.08, 0.28)

Note: OR = Odds Ratio.
***$p < .001$, **$p < .01$, *$p < .05$

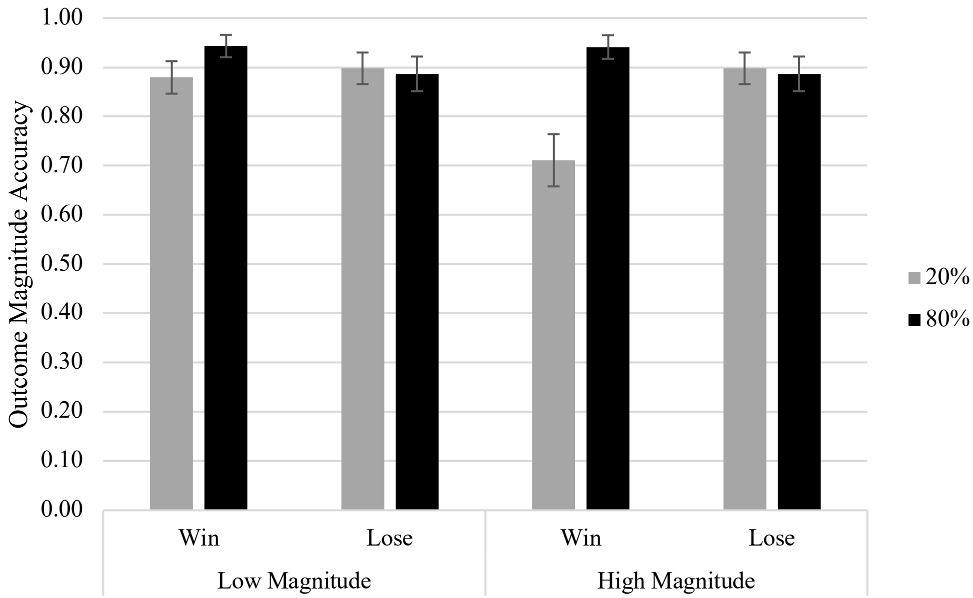


Figure 4.14. Outcome magnitude accuracy.
Accuracy was evaluated for scenes that participants correctly indicated appeared in the VLT.

Correlation Between Value Learning and Memory Accuracy for Win and Loss Scenes

We also examined the relationship between value learning performance for win and loss scenes and accuracy for the corresponding win and loss scenes. A correlation for each image (80% win, 20% win, 80% loss, 20% loss) in each outcome magnitude condition was computed and the significance was determined after correcting for multiple correlations using a Holm-Bonferroni correction. For the low magnitude pairs, no significant correlations emerged between win learning in the VLT and memory accuracy for either win scene (20%: $r(67) = -.03$, $p = .78$; /80%: $r(67) = .28$, $p = .02$). No significant correlations emerged between loss learning for low-magnitude pairs in the VLT and memory accuracy for either loss scene (20%: $r(67) = .28$, $p = .02$/80%: $r(67) = .22$, $p = .07$). For the high magnitude pairs, better win learning in the VLT was correlated with lower memory accuracy for the 20% win scene ($r(67) = -.47$, $p < .001$) but not the 80% win scene, $r(67) = .27$, $p = .03$. No significant correlations emerged between loss learning for high-magnitude pairs in the VLT and memory accuracy for either loss scene (20%: $r(67) = .17$, $p = .16$/80%: $r(67) = .24$, $p = .05$). With the exception of the 20% win scene in the high-magnitude condition, there was no association between level of learning in block 5 of the VLT and accuracy of the outcome associated with each corresponding scene in the memory task. This general pattern is consistent in previous chapters and our previous work (Lin et al. 2020). Therefore, these correlations will not be discussed any further.

Reinforcement Learning

To examine whether the difference in learning for win and loss associations was correlated with the weight that participants placed on old versus new information and balance between exploration and exploitation, a series of correlations were conducted between the learning difference score for each outcome magnitude condition and the corresponding learning rate and inverse temperature parameter. The *p*-value cutoff for significance was adjusted to account for the number of correlations (α/number of correlations).

The results revealed that a smaller learning asymmetry for the low magnitude pairs (i.e., smaller difference score) was associated with more updating (i.e., higher alpha level), $r(69) = -.31$, $p = .01$, and more exploration (i.e., smaller beta value), $r(69) = .43$, $p < .001$ (Figure 4.15).

The same pattern emerged for the high magnitude pairs: a smaller learning asymmetry was associated with more updating, $r(69) = -.27$, $p = .02$, and more exploration, $r(69) = .33$, $p = .005$.

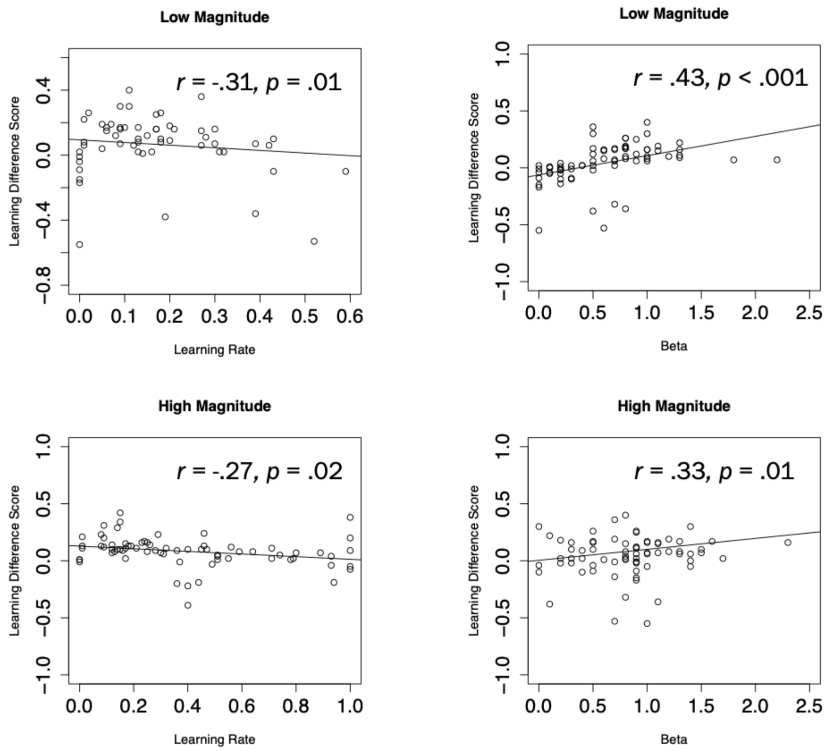


Figure 4.15. Scatterplots of learning difference score and reinforcement learning parameters.

We also examined whether the measure of sensitivity to punishment and reward and loss aversion were associated with the reinforcement learning parameters. A Holm-Bonferroni method was used to adjust the alpha level for multiple comparisons. The results showed no significant correlations between the questionnaire measures and the learning rate and inverse temperature parameter for each outcome magnitude condition (Table 4.8).

Table 4.8.
Pearson r Values for the Correlations between Questionnaire Measures and RL Parameters.

	α -Low Magnitude	β - Low Magnitude	α -Low Magnitude	β - Low Magnitude
Sensitivity to Punishment	.12	.11	.27	-.06
Sensitivity to Reward	-.00	-.14	-.08	-.03
Loss Aversion	.08	.20	.02	.11

Note. α = Learning rate, β = inverse temperature parameter.

General Discussion

Asymmetries in how we respond to gains versus losses has received considerable attention in decision-making research. One of the more well-known observations is that people tend to avoid losses compared to acquiring equivalent gains. In other words, people have a tendency to be loss averse (Kahneman & Tversky, 1979). While the observation that people are loss averse is relatively robust, studies have revealed boundary conditions for loss aversion such that the asymmetric pattern for how people respond to losses versus gains is either reduced or reversed. One example is that when the magnitude of the win and loss outcomes is high, "gains loom larger than losses" (Harinck et al. 2007). In other words, the asymmetry is reversed. This finding demonstrates that the magnitude of outcomes can affect whether people show an asymmetric response towards wins versus losses. In a recent investigation, a learning asymmetry was reported revealing that when it comes to learning to associate neutral stimuli with win and loss outcomes, people learned win associations better than loss associations (Lin et al., 2020). This asymmetry runs in the opposing direction to what would be expected based on the evidence for loss aversion. However, this pattern is consistent with the finding that for small magnitude outcomes, "gains loom larger than losses." This raises the question of whether the learning asymmetry is magnitude dependent. In other words, would increasing the magnitude of the outcomes reduce or reverse the pattern of the learning asymmetry.

The goal of the present study was to test the hypothesis that the magnitude of win and loss outcomes would influence the learning pattern for win and loss associations. We first administered the VLT and memory task to an online sample to establish that the learning asymmetry replicated in an online sample (Experiment 1). Then, in Experiment 2, we varied the magnitude of win and loss outcomes. One win and one loss pair led to small magnitude outcomes (+/- 5 points), whereas another win and loss pair led to large magnitude outcomes (+/- 50 points). We examined the effect of outcome magnitude on learning of win and loss associations, explicit knowledge of the outcomes associated with each image using a post-learning memory task, and evaluated whether trait sensitivity to reward and punishment and loss aversion were correlated with the learning difference between win and loss associations.

Value Learning Task

First, Experiment 1 replicated the learning asymmetry reported by Lin et al. (2020). Participants showed better learning of the optimal choice for the win pair compared with the loss pair. This finding demonstrates the robustness of the learning asymmetry, which to our knowledge, is the first time it is reported for an online sample of participants. Furthermore, our previous studies have included college samples, whereas Experiment 1 included adults with a wider age range residing across the country. Therefore, the sample was more representative of a younger adult population. While the results of Experiment 1 replicated those reported by Lin et al. (2020), one difference between the sets of experiments is the exclusion rate. In the present experiment, 38% of the participants were excluded due to failure to reach the learning criterion. This exclusion rate is higher than what was reported in both experiments conducted by Lin et al. (2020): 18.5% of participants were excluded in Experiment 1, whereas 19.68% were excluded in

Experiment 2. The higher exclusion rate may be an artifact of conducting studies online where the environment is not as controlled as a laboratory setting.

In Experiment 2, several key findings emerged when examining performance on the VLT. First, as in our previous work, participants showed increased learning throughout all trial blocks. Second, participants showed better initial learning for high magnitude pairs versus low magnitude pairs that was evident during earlier blocks but reached a similar level of learning by the end of the task. Third, overall, participants showed better learning for win pairs versus loss pairs. Exploratory analyses revealed that for the low magnitude condition, this learning asymmetry was only present during Trial Block 4. Thus, by the end of the task, participants showed similar learning for wins and losses of small magnitude. When examining learning for the high magnitude condition, participants showed better learning for the win pair versus loss pair throughout all trial blocks. Although it appears that the learning asymmetry was reduced, at least quantitatively, during the last block of trials. These findings are the opposite effect than we predicted. While we hypothesized that the learning asymmetry would be reduced with higher magnitude payoffs, the results showed that the learning asymmetry was reduced for the pairs with lower magnitude payoffs. Nonetheless, the results demonstrate that outcome magnitude does matter in the sense that when multiple magnitude conditions are included in the same context, the learning pattern for win and loss associations varies for low and high magnitude outcomes.

These findings do not align with the prediction that for small magnitude outcomes "gains loom larger than losses," whereas for high magnitude outcomes "losses loom larger than gains" (Harinck et al. 2007). Instead, it appears that participants prioritized learning the high-magnitude win pair in order to maximize positive outcomes, while discounting high magnitude losses and low magnitude wins and losses, which is consistent with the hedonic principle (Higgins, 1997; Kahneman, Diener, & Schwarz, 1999). This led to a reduction in the learning asymmetry in the low magnitude condition and contributed to the learning asymmetry in the high magnitude condition. One potential explanation for why the pattern of results in the present study differ from the predictions made by Harinck et al. (2007) may be the nature of the tasks. In the work reported by Harinck et al. (2007), the effect of outcome magnitude on gain and loss asymmetries was demonstrated in a series of experiments concerning participants' *predictions* of how they would react to hypothetical losses and gains. However, previous research has demonstrated that individuals overestimate the intensity and duration of their emotional reactions to future events compared to what they actually experience (Gilbert, Morewedge, Risen, & Wilson, 2004; Wilson & Gilbert, 2003). In the present study, participants engaged in a task where they actually experienced gains and losses that translated into a monetary outcome. Thus, it is possible that the effect of outcome magnitude on gain and loss asymmetries depends on whether the gains and losses are anticipated or experienced.

Memory Task

The results from Experiment 1 also replicated the pattern of explicit knowledge reported by Lin et al. (2020). Participants were more accurate at identifying the outcome associated with the optimal win scene (80% win scene) versus the outcome associated with the optimal loss scene (20% loss scene). This pattern is expected given that participants sample the optimal win scene more frequently than the optimal loss scene. When comparing explicit knowledge between the suboptimal win scenes and suboptimal loss scenes, participants were more accurate at identifying the outcome associated with the suboptimal loss scenes versus suboptimal win scene, again consistent with their experiences of outcomes due to their choices in the VLT.

In Experiment 2, the results from the post-learning memory task showed several findings. First, when examining explicit knowledge of the outcomes associated with each VLT scene, outcome association accuracy was higher for scenes that led to high magnitude outcomes versus low magnitude outcomes. This finding suggests that participants discounted or devalued learning of the low outcome images. Second, consistent with previous work by Lin et al. (2020) participants were more accurate at identifying the outcome associated with the optimal win scene than the optimal loss scene. This pattern emerged for both magnitude conditions despite the learning asymmetry between win and loss associations diminishing for the low magnitude pairs. Third, for win scenes, accuracy was higher for the optimal win scene versus suboptimal win scene, which is consistent with choice behavior in the VLT for the win pair. Whereas for loss scenes, accuracy did not differ between optimal and suboptimal scenes. When examining misattributions, participants were less likely to make valence misattributions for high probability win scenes and high probability loss scenes when compared with low probability win and loss scenes. Thus, despite a reduced learning asymmetry for wins and losses, explicit knowledge of the outcomes still differed between win and loss scenes for both magnitude conditions. Although overall outcome association accuracy was higher for high-magnitude win scenes versus low-magnitude win scenes, the remaining pattern of results were similar between the low and high-magnitude scenes.

We also examined explicit knowledge of the magnitude conditions. If participants correctly indicated that a scene appeared in the VLT, they were prompted to indicate whether the scene was in the low or high magnitude condition. The pattern for outcome magnitude accuracy mirrored the pattern for outcome association accuracy. For optimal scenes, accuracy was higher for the win scenes versus loss scenes. Whereas for suboptimal scenes, accuracy was higher for loss scenes versus win scenes. Furthermore, for win scenes, accuracy was higher for optimal scenes versus suboptimal scenes. Whereas for loss scenes, accuracy did not significantly differ between optimal and suboptimal scenes. Overall, participants were more accurate at indicating the magnitude condition for low magnitude scenes. This finding was primarily driven by lower outcome magnitude accuracy for the high-magnitude suboptimal win scene, which is not surprising given that the high-magnitude suboptimal win scene was sampled the least out of all the scenes during the VLT.

The pattern of results showing that participants had better explicit knowledge for the optimal win scene versus optimal loss scene, together with the results from the VLT showing that optimal choice was higher for the win pair versus loss pair in the high magnitude condition suggests that learners prioritized learning the wins for the high magnitude condition during the earlier trials. However, as they progressed with the task, they continued learning the optimal choice for the remaining scenes as evidenced by similar rate of optimal selection for all four pairs by the end of the task.

Reinforcement Learning

When examining how the reinforcement learning parameters correlated with the learning difference between wins and losses, the results showed that a more aggressive learning rate (i.e., more updating from recent trials) and more exploration was correlated with a smaller learning difference. Finding that a more aggressive learning update rule is associated with a smaller learning difference is consistent with our previous work (Hao et al., 2019). However, the present study indicated that greater exploration is associated with a smaller learning asymmetry, which differs from previous work suggesting that greater *exploitation* is the optimal choice strategy. One potential explanation for this difference is that in the present study, participants adopted a different choice strategy due to outcomes having higher stakes and there being more scenes to learn. As such, it was more beneficial for participants to explore the outcomes of each scene in order to learn the optimal image for each pair.

Limitations and Future Directions

One limitation of the present study is that the sample size in Experiment 2 may have restricted the statistical power to detect individual differences in sensitivity to reward and punishment and loss aversion. The sample size was substantially reduced after excluding participants that failed to reach the learning criterion (58% were excluded in Experiment 2). This exclusion rate was higher compared to our previous studies reported in Lin et al. (2020). The reason(s) for this higher exclusion rate are unclear. One possibility may have been due to the nature of online data collection where the setting is less controlled compared to what is typically controlled for during in-person data collection, such as minimizing external distractions.

In the present study, we included a performance-based incentive to make the different outcome magnitude conditions more salient to participants. However, our previous work provided evidence for the equivalence between points and monetary incentives based on the finding that the learning asymmetry was not reduced when participants were provided with a monetary incentive (Lin et al. 2020). Whether this equivalence generalizes to a magnitude manipulation remains an open question for future exploration.

In conclusion, the present work provides evidence that including multiple outcome magnitude conditions in the same context affects the learning pattern for win and loss associations. While the learning asymmetry emerged for pairs with high magnitude payoffs, the asymmetry was reduced for pairs with lower magnitude payoffs. We propose that this finding can be explained by the observation that people have a tendency to focus on maximizing positive outcomes while minimizing negative outcomes (i.e., hedonic principle). Thus, participants discounted high-magnitude losses and small magnitude payoffs. These results build on the framework identifying boundary conditions for gain and loss asymmetries. While previous studies have noted the effect of outcome magnitude on predictions of how people will react to losses and gains, we provide evidence that different outcome magnitude conditions in the same context can also affect participants actions in approaching wins and avoiding losses. Furthermore, the effect of outcome magnitude on gain and loss asymmetries may differ for people's *predictions* of how they would react to hypothetical losses and gains and choice behavior for maximizing gains and minimizing losses.

References

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5,* 323–370. http://dx.doi.org/10.1037/1089-2680.5.4.323

Erev, I., Ert, E., & Yechiam, E. (2008). Loss aversion, diminishing sensitivity, and the effect of experience on repeated decisions. *Journal of Behavioral Decision Making, 21*(5), 575–597. https://doi.org/10.1002/bdm.602

Ert, E. & Erev, I. (2013). On the descriptive value of loss aversion in decisions under risk: Six clarifications. *Judgment and Decision Making, 8*(3), 214-235. http://dx.doi.org/10.2139/ssrn.1012022

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160.

Gilbert, D.T., Morewedge, C.K., Risen, J.L., & Wilson, T.D. (2004). Looking forward to looking backward: The misprediction of regret. *Psychological Science, 15*, 346–350.

Harinck, F., Van Dijk, E., Van Beest, I., & Mersmann, P. (2007). When gains loom larger than losses: Reversed loss aversion for small amounts of money. Psychological Science, 18(12), 1099–1105. https://doi.org/10.1111/j.1467-9280.2007.02031.x

Higgins, E.T. (1997). Beyond pleasure and pain. *American Psychologist, 52*, 1280–1300.

Kahneman, D., Diener, E., & Schwarz, N. (1999). Well-being: The foundations of hedonic psychology. New York: Russell Sage Foundation.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk.

Econometrica, 47, 263–291.

Kahneman, D., Knetsch, Jack & Thaler, Richard. (1990). Experimental tests of the endowment effect and the coase theorem. *Journal of Political Economy 98*, 1325-1348. Doi.10.1086/261737.

Kermer, D.A., Driver-Linn, E., Wilson, T.D., & Gilbert, D.T. (2006). Loss aversion is an affective forecasting error. *Psychological Science, 17*, 649–653.

Lin, Z., Cabrera-Haro, L. E., & Reuter-Lorenz, P. A. (2020). Asymmetrical learning and memory for acquired gain versus loss associations. Cognition, *202*:104318. https://doi.org/10.1016/j.cognition.2020.104318

O'Brien, J. L., & Raymond, J. E. (2012). Learned predictiveness speeds visual processing. *Psychological Science, 23*(3), 359-363. doi: 10.1177/0956797611429800

Painter, D. R., Kritikos, A., & Raymond, J. E. (2014) Value learning modulates goal-directed actions. *The Quarterly Journal of Experimental*

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods.* 10.3758/s13428-018-01193-y

Raymond, J. E. & O'Brien, J. L. (2009). Selective visual attention and motivation: The consequences of value learning in an attentional blink task. *Psychological Science, 20*(8), 981-988. doi: 10.1111/j.1467-9280.2009.02391.x

Rissman, J., Gazzaley, A., & D'Esposito, M. (2009). The effect of non-visual working memory load on top-down modulation of visual processing. *Neuropsychologia, 47*(7), 1637-1646.

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion.

Personality and Social Psychology Review, 5, 296–320. http://dx.doi.org/10.1207/S15327957PSPR0504_2

Stewart, N., Chater, N., & Brown, G.D.A. (2006). Decision by sampling. *Cognitive Psychology, 53*(1), 1–26.

Torrubia, R., Ávila, C., Moltó, J., & Caseras, X. (2001). The sensitivity to punishment and sensitivity to reward questionnaire (SPSRQ) as a measure of Gray's anxiety and I mpulsivity dimensions. *Personality and Individual Differences, 31*(6), 837-862.

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin, 134*(3), 383-403. doi: 10.1037/0033-2909.134.3.383

Wilson, T. D., & Gilbert, D. T. (2003). Affective forecasting. *Advances in Experimental Social Psychology, 35*, 345– 411. http://dx.doi.org/10.1016/S0065-2601(03)01006-2

Supplemental Material

The analyses presented in Experiment 2 were repeated for participants who did not reach the learning criterion (non-learners; 97 participants).

Value Learning Task

To evaluate performance on the VLT, a 5 (Trial Block: 1-5) x 2 (Outcome Valence: win versus loss) x 2 (Outcome Magnitude: low versus high) within-subjects ANOVA was conducted with probability of optimal choice as the dependent variable. When evaluating main effects and interactions, Greenhouse-Geisser corrections were applied when assumptions of sphericity were violated. The *p*-value significance cutoffs were adjusted for follow-up pairwise comparisons to adjust for multiple comparisons using a Bonferroni correction (α/number of comparisons).

The results showed a significant main effect of Outcome Magnitude indicating that probability of optimal choice was higher for scene pairs that led to outcomes of high magnitude ($M = .59$, $SE = .02$) versus low magnitude outcomes ($M = .55$, $SE = .01$), $F(1, 96) = 5.00$, $p = .03$, $n_{\mathrm{p}}^2 = .05$ (Figure 4.16). A significant main effect of Trial Block also emerged, $F(2.55, 244.50) = 10.32$, $p < .001$, $n_{\mathrm{p}}^2 = .10$. Follow-up pairwise comparisons revealed that probability of optimal choice was higher for Trial Block 4 when compared with Trial Blocks 1 and 2, *p*s < .009, and was higher for Trial Block 5 when compared with Trial Blocks 1 and 2, *p*s < .002. Thus, learning improved by the end of the task. No significant main effect emerged for Outcome Valence, suggesting that probability of optimal choice did not significantly differ between the win ($M = .56$, $SE = .01$) and loss pairs ($M = .57$, $SE = .01$), $F(1, 96) = 0.65$, $p = .42$.

When evaluating the two-way interactions, a significant interaction emerged between Outcome Magnitude x Trial Block, $F(2.96, 284.04) = 2.92$, $p = .04$. Follow-up pairwise comparisons revealed that probability of optimal choice was higher for scene pairs that led to high magnitude outcomes versus low magnitude outcomes only during Trial Block 3, $p = .006$. No significant differences emerged for the remaining Trial Blocks. No significant interactions emerged for Valence x Outcome Magnitude, $F(1, 96) = 0.07$, $p = .79$ and Valence x Trial Block, $F(2.65, 253.97) = 1.61$, $p = .19$. Lastly, the three-way interaction between Valence x Outcome Magnitude x Trial Block was not significant, $F(3.09, 296.62) = 1.37$, $p = .25$.

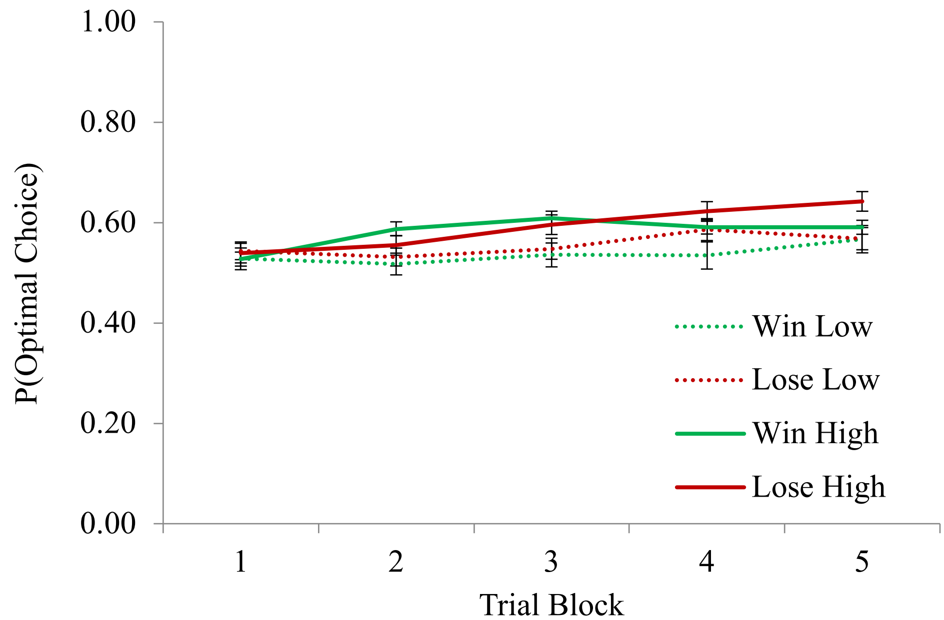


Figure 4.16. *Performance on the VLT for participants who did not meet the learning criterion.*

Correlations between Questionnaire Measures and Performance on the VLT

We also examined whether measures of sensitivity to punishment and reward and loss aversion were associated with the learning difference between win and loss outcome associations for non-learners. Similar to the results for the learners, no significant correlations emerged, $ps > .14$ (Table 4.9).

Table 4.9.
Correlational Analyses between the Learning Difference Score and Measures of Sensitivity to Punishment and Reward and Loss Aversion.

Learning Difference	Sensitivity to Punishment	Sensitivity to Reward	Loss Aversion
Low magnitude pairs	.00	.15	-.03
High magnitude pairs	.14	.01	-.00

Note. The learning difference score was calculated for each magnitude condition as p(optimal choice) averaged across all trials for the win pair minus loss pair. *p*-value cutoffs were adjusted to account for the number of correlations (α/number of correlations).
***$p < .001$, **$p < .01$, *$p < .05$

Memory Task

Memory accuracy for VLT Scenes. As in the analysis for learners, we used a binary logistic regression analysis to examine explicit knowledge of the outcomes associated with each VLT scene. Valence (win versus loss) Optimality (suboptimal versus optimal), and Outcome Magnitude (low versus high) were entered as predictors of outcome accuracy. Interaction terms were added iteratively to examine whether each term increased model fit. Interaction terms which significantly increased model fit were included in the final model. Thus, the final model included Valence, Optimality, Magnitude, and Optimality x Outcome Magnitude as predictors of outcome accuracy (Table 4.10).

The results showed that Optimality was a significant predictor of outcome accuracy. Accuracy was higher for optimal scenes versus suboptimal scenes. However, the Optimality x Outcome Magnitude interaction revealed that this pattern only held for scenes that led to low magnitude outcomes (Figure 4.17). For scenes that led to high magnitude outcomes, no significant difference in accuracy emerged between optimal and suboptimal scenes. Outcome Magnitude was also a significant predictor of outcome accuracy. Accuracy was higher for scenes that led to high magnitude outcomes versus low magnitude outcomes. The Optimality x Outcome

Magnitude interaction revealed that this pattern held for both optimal and suboptimal scenes but was more pronounced for suboptimal scenes.

Table 4.10.
Binary Logistic Regression Predicting Memory Accuracy from Valence, Optimality, and Outcome Magnitude.

Predictor	β	OR	95% CI for OR
Intercept	-1.86***	0.16	(0.12, 0.20)
Valence	-0.10	0.90	(0.76, 1.07)
Optimality	0.67***	1.95	(1.49, 2.55)
Outcome Magnitude	0.97***	2.66	(2.05, 3.46)
Optimality*Outcome Magnitude	-0.65***	0.52	(0.37, 0.74)

Note: OR = Odds Ratio. Predictors were coded as follows: Valence (win coded as 0), Optimality (suboptimal coded as 0), Outcome Magnitude (low magnitude coded as 0). For outcome accuracy, incorrect outcome coded as 0, correct outcome was coded as 1.
***p < .001, **p < .01, *p < .05

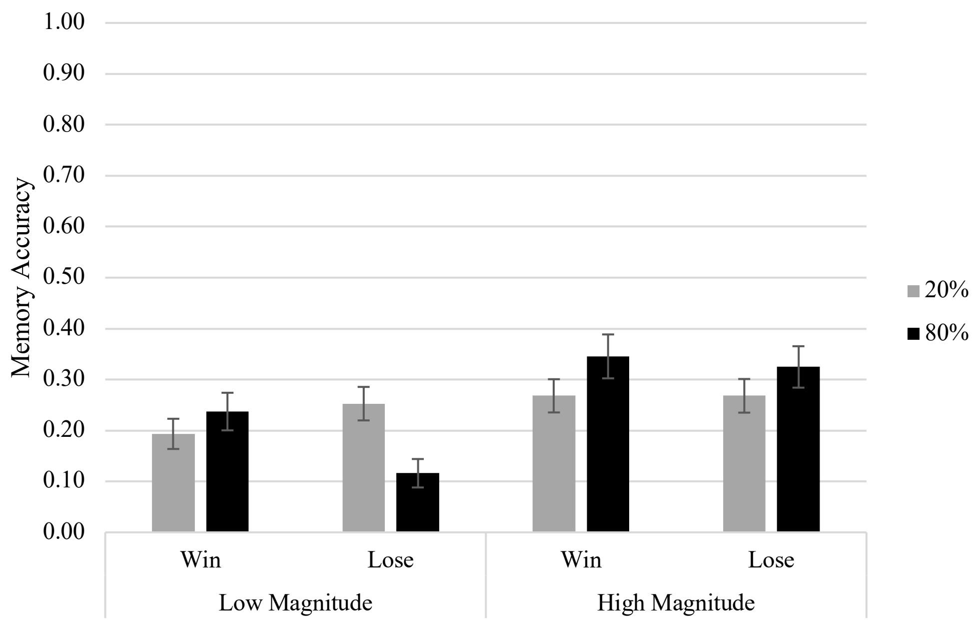


Figure 4.17. Accuracy of explicit knowledge of outcomes associated with the VLT scenes.

Outcome Attributions. We also examined the full range of participants' responses on the Memory Task for each VLT scene using a multinomial logistic regression. Valence, Probability, and Outcome Magnitude were predictor variables, while the probability that each response option was selected was the outcome variable (Table 4.11).

Outcome attributions for win pairs. For the 80% win scene in the low magnitude condition, participants were more likely to attribute "very likely to win" and "occasionally win" as the outcome associated with the scene (Figure 4.18). The remaining responses were distributed similarly across "no change," "occasionally lose," and "none(new)." For the 80% in the high magnitude condition, participants selected "very likely to win," more often than other response options. The remaining response options were primarily "occasionally win" and "no change." Thus, for the 80% win scenes, while attribution errors included confusion in the probability of a win and valence misattribution, participants were less likely to attribute a high probability loss to the 80% win scene.

For the 20% win scene in the low magnitude condition, participants were more likely to select "no change," followed by "occasionally win," "occasionally lose," and "none(new)." For the 20% win scene in the high magnitude condition, participants were more likely to select "occasionally win" and "no change." Other attributions were distributed similarly across the remaining response options. Thus, for the 20% win scenes, attribution errors included confusion in the probability of a win and valence misattributions.

Outcome attributions for loss pairs. For the 80% loss scene in the low magnitude condition, participants were more likely to attribute "occasionally lose" as the outcome associated with the scene. Other responses were primarily "occasionally win" and "no change." For the 80% loss scene in the high magnitude condition, participants were more likely to attribute “very likely to lose” as the outcome associated with the scene. Other responses were primarily “occasionally win,” “no change,” and “none(new).” Thus, while attribution errors for the 80% loss scenes, included confusion in the probability of a loss and valence misattribution, participants were less likely to attribute a high probability win to the 80% loss scene.

For the 20% loss scene in the low magnitude and high magnitude conditions, participants were just as likely to select “occasionally lose,” “no change,” and “occasionally win” as the outcome associated with the scene. Other attributions for the 20% loss scene in the low magnitude condition were primarily “none(new),” while attributions for the 20% loss scene in the high magnitude condition included “very likely to lose” and “none(new).” Thus, for the 20% loss scene, attribution errors included valence misattributions but were less likely to misattribute a high probability loss or win to the 20% loss scene.

In summary, these results revealed that non-learner’s misattributions for wins and losses included confusion in the probability of the valenced outcome and valence misattributions. This pattern emerged for scenes that led to low and high magnitude outcomes.

Table 4.11.
Multinomial Logistic Regression Predicting Outcome Response for each VLT scene from Valence and Probability.

Outcome Response	Valence	Probability	Outcome Magnitude	Valence * Probability	Valence * Magnitude	Probability * Magnitude	Valence * Probability * Magnitude
Very likely to win	0.83	6.27***	1.91*	0.23**	0.60	0.70	1.97
Occasionally win	1.29	2.37***	1.83**	0.64	0.55	0.39**	2.04
Occasionally lose	1.39	1.66*	1.09	1.41	1.03	0.59	0.89
Very likely to lose	0.68	1.36	1.72*	2.13	1.19	0.34*	5.24**
None (new)	0.86	1.55	1.09	0.78	0.85	0.80	1.88

Note. Values in the table are odds ratio. Valence was coded as 0 for the win scenes, and 1 for the loss scenes. Probability was coded as 0 for low probability, and 1 for high probability. Outcome magnitude was coded as 0 for low magnitude, and 1 for high magnitude.

*$p < .05$, ** $p < .01$, *** $p < .001$

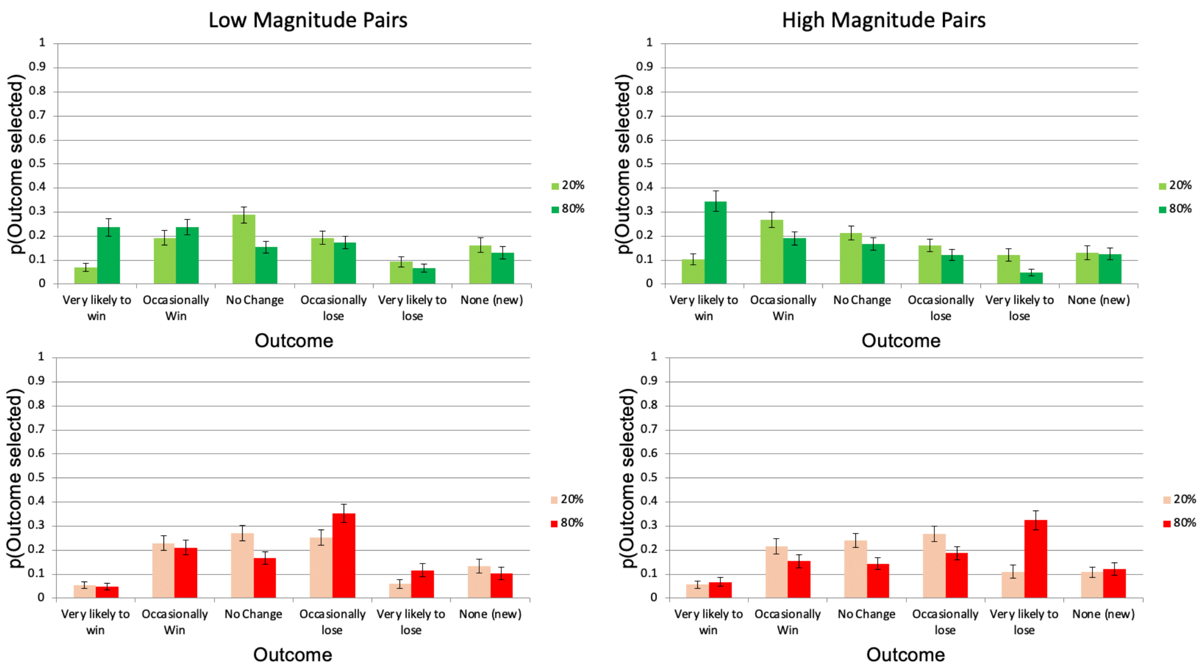


Figure 4.18. Outcome attributions in the Memory Task.
For each VLT scene, participants were instructed to indicate the outcome most likely associated with the scene.

Explicit knowledge of outcome magnitude. We also examined accuracy in attributing the outcome magnitude of scenes for which participants correctly indicated appeared in the VLT. We used a binary logistic regression analysis with Valence (win versus loss), Optimality (suboptimal versus optimal), and Outcome Magnitude (low versus high) as predictors of outcome magnitude accuracy. Interaction terms were added iteratively to examine whether each term increased model fit. None of the interaction terms significantly increased model fit. Therefore, the final model included Valence, Optimality, and Magnitude as predictors of outcome magnitude accuracy.

The results revealed that Outcome Magnitude was the only significant predictor of outcome magnitude accuracy (Table 4.12). Participants were more accurate at identifying the magnitude associated with low magnitude scenes versus high magnitude scenes (Figure 4.19).

Table 4.12.
Binary Logistic Regression Predicting Accuracy of Outcome Magnitude from Valence, Optimality, and Outcome Magnitude.

Predictor	β	OR	95% CI for OR
Intercept	0.91***	2.49	(1.89, 3.30)
Valence	-0.01	0.99	(0.84, 1.18)
Optimality	-0.08	0.92	(0.78, 1.09)
Outcome Magnitude	-0.40***	0.67	(0.56, 0.79)

Note: OR = Odds Ratio.
****p* < .001, ***p* < .01, **p* < .05

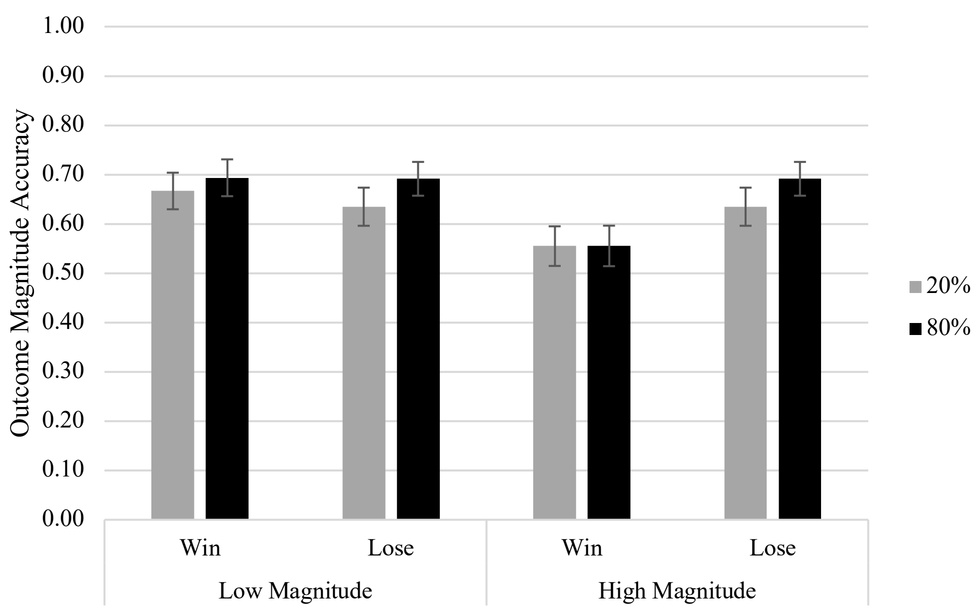


Figure 4.19. Outcome magnitude accuracy.

Reinforcement Learning

To examine whether the difference in learning for win and loss associations was correlated with the weight that participants placed on old versus new information and balance between exploration and exploitation, a series of correlations were conducted between the learning difference score for each outcome magnitude condition and the corresponding learning rate and inverse temperature parameter. The *p*-value cutoff for significance was adjusted to account for the number of correlations (α/number of correlations).

For non-learners, the learning difference score for low magnitude pairs was not significantly correlated with the learning rate, $r(95) = -.17$, $p = .09$, or inverse temperature parameter, $r(95) = .21$, $p = .04$ (Figure 4.20). The learning difference score for high magnitude pairs was also not significantly correlated with the learning rate, $r(95) = -.03$, $p = .74$, or inverse temperature parameter, $r(95) = .21$, $p = .04$.

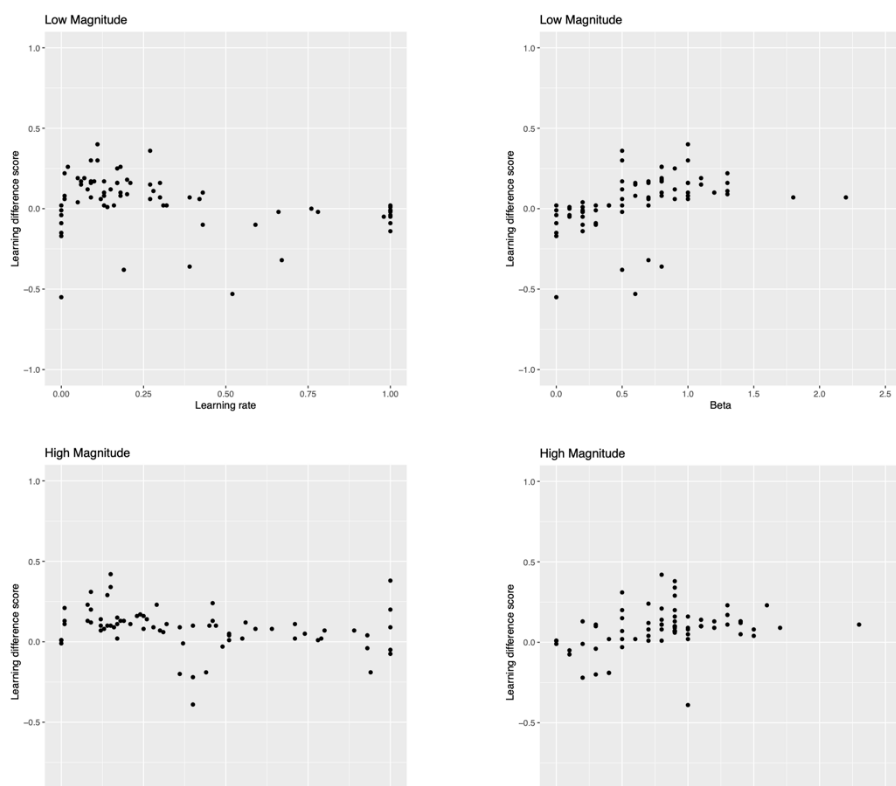


Figure 4.20. Scatterplots of learning difference score and reinforcement learning parameters.

Chapter 5

General Discussion

When it comes to allocating our attention toward information in our environment, previously learned associations between stimuli and positive or negative outcomes (i.e., value) can influence what we attend to or ignore. A number of studies have examined how learned value can influence different cognitive processes, such as attention, perceptual processing, and motor control (O'Brien & Raymond, 2012; Painter, Kritikos, & Raymond, 2013; Raymond & O'Brien, 2009; Rutherford, O'Brien & Raymond, 2010). These studies have reported differential effects of learned value on subsequent processing of originally neutral stimuli that have acquired value through trial-and-error learning. However, the extent to which initial learning of win and loss associations is equivalent has been relatively unexplored. Recently, a learning asymmetry was revealed for a probabilistic Value Learning Task (VLT), such that win associations were learned better than loss associations (Lin et al., 2020). This learning asymmetry poses a potential challenge for conclusions drawn about the effects of learned value on subsequent stimulus processing as measured in cognitive and motor tasks. That is, learning differences between win and loss associations could be the basis for valence effects in such tasks rather than differences in value per se.

Despite the robust advantage for learning win versus loss associations demonstrated by Lin et al. (2020), questions about the boundary conditions for this learning asymmetry remained unanswered. Thus, this dissertation focused on several potential boundary conditions for the learning advantage favoring wins in order to understand if and why the learning asymmetry emerges under some conditions and not others. The aims of the present dissertation were as follows: 1) to investigate whether the learning asymmetry previously reported by Lin et al. (2020) is a dimension along which individuals vary and whether learning patterns are consistent within individuals; 2) to investigate whether the learning asymmetry depends on the context in which win and loss associations are learned; and 3) to investigate whether the magnitude of outcomes affects the learning pattern for win and loss associations. A summary of the key findings is provided below, followed by a discussion of their implications.

Summary of Key Findings

Learning patterns for win and loss associations. The first aim of this dissertation was to investigate whether the learning advantage for win associations compared to loss associations is universal and a fundamental property of the VLT, or whether there are individual difference variables that influence this pattern. In Chapter 2, this was accomplished by having participants complete the VLT on two separate occasions. We tested the hypothesis that there are underlying individual differences in the extent to which win associations are learned better than losses and examined whether patterns of asymmetrical learning were consistent within an individual across testing sessions. The results showed that across the entire sample of participants, the learning asymmetry favoring win associations was prevalent during both sessions, although it appeared quantitatively less robust in the second session. When examining consistency in learning, the learning difference between wins and losses (i.e., an index of asymmetry magnitude) for Session 1 and 2 were not correlated. However, when evaluating win and loss learning separately, loss learning, but not win learning, was consistent across sessions. One potential explanation for the lack of a significant correlation for win learning is a lack of variability in win learning, given that win learning tends to be at ceiling. Indeed, when we compared the variances for win and loss learning, win learning had lower variance. We also found no significant correlations between the learning difference for wins and losses and individual difference measures of (1) behavioral approach and avoidance, (2) impulsivity, or (3) sensitivity to reward and punishment. Thus, there was no strong evidence for intraindividual consistency in learning patterns (with the exception of loss learning) or evidence that learning patterns were associated with traits such as impulsivity, or sensitivity to reward and punishment. The finding that loss learning for Session 1 and 2 were positively correlated suggests that participants differ in the extent to which they learn the optimal image for avoiding losses. That is, participants who had better loss learning during Session 1 also had better loss learning during Session 2. One possible explanation for the finding that participants were consistent in their loss learning is that those who consistently had better loss learning were those who incorporated strategies to learn the task structure, which they incorporated during Session 2.

The second aim of this dissertation was to investigate whether the learning asymmetry depends on the context in which win and loss associations are learned. Previous studies of the VLT present the win, loss, and no change trials in a randomly interleaved format which requires that win and loss associations are learned concurrently. The experiment reported in Chapter 3 systematically varied the context in which win and loss trials appeared. The win, loss, and no change trials were presented in different trial blocks allowing participants to learn each pair-outcome association separately. Blocking also entailed different valence order conditions, and thus the effect of valence order on learning of win and loss associations was also examined. Consistent with our hypothesis, the results showed that the context in which win and loss associations are presented affects the degree, and indeed, the presence or absence of the learning asymmetry. Participants in the conditions Lose-Win-No Change or No-Change-Lose-Win learned the win pair better than the loss pair. Participants who viewed the pairs in the remaining valence order conditions (Win-Lose-No Change, Lose-No Change-Win, No Change-Win-Lose, Win-No Change-Lose) had similar learning for the win and loss pair. Importantly, the learning asymmetry was attributed to poorer loss learning rather than fluctuations in the level of win learning. Thus, it appears that learning for win associations was relatively stable, whereas loss learning was disadvantaged when loss learning blocks directly preceded win learning blocks. Loss learning, therefore, seems more sensitive than win learning to contextual effects.

The third aim of this dissertation was to investigate whether the magnitude of outcomes affects the learning pattern for win and loss associations. Previous work has demonstrated that the magnitude of wins and losses affects whether loss outcomes have larger subjective weight than symmetric gain outcomes (i.e., loss aversion; Harinck et al. 2007; Kahneman & Tversky, 1979). In our previous work, the learning advantage for wins appeared when earnings (either point or monetary) were relatively small. Therefore, we hypothesized that increasing the magnitude of win and loss outcomes would reduce the learning asymmetry. In Chapter 4, we tested this hypothesis by varying the magnitude of win and loss outcomes. One win and one loss pair led to low magnitude outcomes (+/- 5 points—the same payoffs used in our prior studies) while a second win and loss pair led to high magnitude outcomes (+/- 50 points). Overall, learning was better for the high magnitude pairs versus low magnitude pairs, and the asymmetry favoring win learning was present for both magnitude conditions. Contrary to our hypothesis however, the learning advantage for wins was reduced for small magnitude payoffs compared to a numerically larger asymmetry for high magnitude payoffs. Although the interaction of valence and magnitude was not significant, the direction of the magnitude effects was opposite to our prediction. Nonetheless, the results suggest that including multiple outcome magnitude conditions in the same context may influence the learning pattern for win and loss association.

Explicit knowledge of learned value. Across all these studies, we administered a post-learning memory task to examine participant's explicit knowledge of outcomes associated with each VLT scene. In our previous work, the results from the memory task indicated that explicit knowledge of associated outcomes was superior for optimal win scenes versus optimal loss scenes, suggesting an asymmetry in explicit knowledge (Lin et al. 2020). In Chapter 2, participants completed the memory task after learning 6 different scene-outcome pairings over the course of two sessions. However, their explicit knowledge of outcome associations was tested for the three pairs learned in the second session. The pattern of results for the memory task differed from prior pattern in Lin et al. (2020) in that explicit knowledge did *not* differ between wins and losses. However, another way in which the results differed is that performance in the memory task was substantially lower in the experiment reported in Chapter 2 (< 40% accuracy for each VLT scene). It is possible that by administering the memory task during the second session, participants experienced proactive interference. Thus, the conclusions that can be drawn about explicit memory in this study are limited by the overall poor performance in this task.

In Chapter 3 and Chapter 4, the memory task results provided evidence that whether or not there is a learning asymmetry in the VLT, explicit knowledge of the outcomes differs for win and loss scenes. Overall, participants had better memory for the optimal win scene than the suboptimal win scene, which is unsurprising given that participants sampled the optimal win scene more frequently in the VLT. However, for the loss scene, participants had better explicit knowledge of the outcome associated with the image they needed to avoid in order to minimize losses (i.e., suboptimal loss scene versus optimal loss scene)—and yet overall, the suboptimal loss scene was sampled far less frequently than the optimal loss scene. Note however in Chapter 4, for loss scenes, explicit knowledge did not significantly differ between optimal and suboptimal scenes. Overall, these patterns are similar to what was reported by Lin et al. (2020). Taken together, the results suggest that the memory representations for the learned stimulus-outcome associations acquired in the VLT may not be fully accessible to participants for the purpose of explicit report. This would coincide with previous research demonstrating that explicit knowledge of outcome contingencies is not necessary to establish value-driven attentional biases (Anderson, 2015; Theeuwes & Belopolsky, 2012).

These findings also suggest that the memory representation of stimulus-outcome associations depend in part on choice history, even if the end point of learning is similar between wins and losses. One way to understand the role that choice history plays in forming the memory representations of outcome associations is based on the win-stay, lose-shift strategy (Nowak & Sigmund, 1993). According to this strategy, people are more likely to change their response after experiencing a loss, while maintaining the same response option after experiencing a win. Thus, in the context of the VLT, when participants select the optimal win image (80% win scene) for the win pair, they stay with that choice over subsequent win trials. As a result, they rarely select the suboptimal win image (20% win scene) learning very little about the suboptimal win image. This is reflected in poor outcome association accuracy for the suboptimal win image. On the other hand, for the loss pair, if participants have a tendency to shift to the alternative image pair after experiencing a loss, then participants inconsistently select the optimal loss image (20% loss scene) for the loss pair. This allows participants to learn more about both loss images. This possibility is supported by performance in the memory task in that although overall performance is lower for the loss pairs versus win pairs, participants have better explicit knowledge of the suboptimal loss scene when compared to the suboptimal win scene. Thus, choice history and the frequency of feedback may be playing a role, at least in part, in explicit knowledge of stimulus-outcome associations.

Computational characterization of choice behavior in the VLT. In all experiments, we incorporated reinforcement learning theory to derive the learning rate and inverse temperature parameters for each participant. In Chapter 2, we used these parameters to further test whether the learning asymmetry is a dimension along which individuals vary and whether there is intra-individual consistency across sessions. First, we examined whether the learning difference between wins and losses (probability of optimal choice for the win pair – probability of optimal choice for the loss pair) were associated with each parameter. We found that a more aggressive learning rate was associated with a smaller learning difference between wins and losses during Session 1 but no significant correlation emerged for Session 2. One potential reason for the non-significant correlation during the second session is that with the reduction of the learning asymmetry there was less variability in the learning difference score. The inverse temperature parameter was not associated with the learning difference score for Session 1, but greater exploitation was associated with a smaller learning difference between wins and losses at Session 2. Second, we examined the consistency of the parameters across the two sessions. The results showed that participants who had a more aggressive learning update rule during the first session tended to also update the expected value from recent trials during the second session. Furthermore, when comparing the learning rate between the two sessions, the results showed that the learning rate was higher during the second session. This finding, together with the result that a higher learning rate is associated with a smaller learning difference score suggests that some meta-learning may have occurred after the first session, which led participants to adopt a more aggressive learning rate during the second session. On the other hand, participants' balance between exploration versus exploitation was not consistent across the two sessions. This finding suggests that participants adopted a different strategy when it came to exploring different options versus exploiting what they learned. It is possible that this shift in strategy was also influenced by some general knowledge about the task structure that was acquired after the first session.

In Chapter 3, we compared the learning rate and inverse temperature parameters across valence order conditions. The inverse temperature parameter did not differ significantly for any of the valence order conditions. When comparing the learning rate, participants in the No Change-Lose-Win condition (one of the conditions where the learning asymmetry emerged) displayed a lower learning rate compared to participants in No Change-Win-Lose and Win-No Change-Lose condition. In Chapter 4, we examined the correlation between the reinforcement learning parameters and the learning difference between wins and losses. The results showed that more updating from recent trials and more exploration was associated with a smaller learning difference. This pattern held for pairs that led to low and high magnitude outcomes. The finding that more updating from recent trials is associated with a smaller learning difference is consistent with the results from the previous chapters. Whereas the finding that more exploration is associated with a smaller learning difference is unique to the study reported in Chapter 4.

Implications for Boundary Conditions on the Learning Asymmetry

Taken together, the present findings further document the robustness of the learning asymmetry, which persists across multiple sessions, whether win, loss, and no change pairs are presented in a randomly interleaved presentation or in different trial blocks, and with low or high magnitude outcomes. One important observation is that when the learning asymmetry emerges, loss learning is disadvantaged—the opposite asymmetry was not observed under the conditions explored in this dissertation. Nonetheless, the learning context matters. The results from the present dissertation revealed several conditions that influence whether or not a learning asymmetry emerges, therefore suggesting boundaries for the learning asymmetry. In Chapter 2, the learning asymmetry was reduced, at least quantitatively, during the second instance in which participants completed the VLT. We suggest that the reduction of the learning asymmetry can be attributed to some meta-knowledge that participants obtained from the first session, such as how many pairs of scenes to expect, knowledge of three valence conditions (win, loss, and no change), or that each scene within each pair has a different probability of leading to win and loss outcomes. Note that Lin et al. (2020) documented that the magnitude of the learning asymmetry was not affected by explicit instruction about the task structure. Therefore, we surmise that what is learned in the first of our two-session study may have been procedural or implicit knowledge of the task structure or other strategies about trial-and-error learning. Alternatively, task experience may have influenced the relative value or meaning of losses—a possibility we address further below. Future studies will be needed to gain further insight into what knowledge participants acquired from the first session.

In Chapter 3, we demonstrated that the context in which participants learn wins and losses affects the presence or absence of a learning asymmetry. The learning asymmetry only appeared in conditions in which loss learning directly preceded wins. One possibility is that prior experience with wins may make losses more meaningful. That is, once participants experience wins, they may focus on preserving their earnings by learning how to avoid losses. Lastly, in Chapter 4, we demonstrated that including multiple magnitude payoffs affects the learning pattern for win and losses. The learning asymmetry emerged for high magnitude payoffs but was reduced quantitatively for low magnitude payoffs. Thus, these findings build on previous decision-making research demonstrating how asymmetric responses to wins and losses can be altered (Erev, Ert, & Yechiam, 2008; Harinck,van Dijk, Van Beest, & Mersmann, 2007).

Applying reinforcement learning theory to characterize choice behavior in the VLT also provided insight into strategies that are associated with a smaller learning asymmetry. Across all experiments, when a reliable association between the learning rate and the learning difference score emerged, they were inversely related. That is, the higher the learning rate, the smaller the learning asymmetry. This finding is consistent with previous studies suggesting that a more aggressive update rule, by incorporating feedback from recent trials, is beneficial toward learning the optimal choice (Hao et al., 2019). The pattern of results for the inverse temperature parameter were mixed. In Chapter 2, greater exploitation during the second session of completing the VLT was associated with a smaller learning difference. In Chapter 3, the inverse temperature did not differ between valence order conditions. In Chapter 4, for both outcome magnitude conditions, more exploration was associated with a smaller learning difference. One potential explanation for these findings is that the conditions that affect the presence or absence of the learning asymmetry may also influence whether exploring versus exploiting is beneficial for learning the optimal choice. In the experiment reported in Chapter 4, participants had to learn two more valenced pairs compared to the other experiments. Thus, it is possible that more exploration was required for participants to learn the optimal choice across all four pairs.

Implications for Research on Individual Differences in Win and Loss Learning

Previous research has demonstrated that the extent to which people approach positive outcomes more than avoid negative ones (or vice versa), is associated with factors such as particular genetic polymorphisms (e.g., Frank & Hutchison, 2009), different levels of dopamine function (e.g., Frank, Seeberger, & O'Reilly, 2004), and different approach and avoidance personality styles (e.g., Aberg, Doell, & Schwartz, 2016). In the present dissertation, while we showed that participants differed in the extent to which they learned win associations better than loss associations, we found no evidence that the learning difference was associated with measures of trait sensitivity to reward and punishment, behavioral inhibition and avoidance, or impulsivity. Instead, we found that the extent to which participants rely on feedback from recent trials (i.e., learning rate) and balance the degree to which they explore versus exploit learned value is associated with the learning difference between wins and losses. However, these individual tendencies were not consistent across sessions. The only consistencies were that loss learning during Session 1 and 2 were positively correlated and the learning rate for Session 1 and 2 were positively correlated. Thus, other factors may play a role in distinguishing different learning patterns. For example, we suggest that the extent to which meta-knowledge acquired after the first session is incorporated during Session 2 contributes to better loss learning.

Implications for Theories of Decision-Making

The direction of the learning asymmetry is opposite to a major tenet of decision-making. That is, one prominent observation is that loss outcomes have larger subjective weight than symmetric gain outcomes (i.e., loss aversion, Kahneman & Tversky, 1979). Asymmetric effects for losses compared to wins have been demonstrated in both psychological and physiological responses. For example, people have reported greater distress about losing money versus excitement about winning money (McGraw, Larsen, Kahneman, & Schkade, 2010). Losses have been found to have larger effects than gains on physiological arousal, such as pupil diameter, heart rate, and cortical sensitivity (e.g., Hochman & Yechiam, 2011; Tom, Fox, Trepel, & Poldrack, 2007). Therefore, one might expect that learning pertaining to loss outcomes should be more robust than learning for win outcomes. However, in the present work, no conditions favored loss learning relative to win learning, which was robustly advantaged overall. This set of observations provides further evidence that losses do not always loom larger than gains (c.f., Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Rozin & Royzman, 2001; Vaish, Grossmann, & Woodward, 2008).

The results in Chapter 3 build on research demonstrating that encountering information of a given valence can influence subsequent judgments of information of the opposite valence (i.e., framing effects; Ledgerwood & Boydstu, 2014; Sparks & Ledgerwood, 2017). When separating learning for wins and losses, a learning asymmetry only emerged when loss learning directly preceded win learning and loss learning was disadvantaged rather than win learning being advantaged. Furthermore, the results in Chapter 3 suggest a revision of the attentional framework proposed by Yechiam and Hochman (2013). According to their attentional framework, losses drive how attention is invested towards a task and in turn sensitivity towards task reinforcements. Instead, we showed that the temporal order of wins appears to heighten motivation pertaining to potential losses. In the context of value learning, attention and sensitivity toward loss payoffs may vary as a function of whether they appear in the same condition as wins, and the temporal order of wins and losses. Thus, there may be different conditions under which losses drive attention toward wins or vice versa.

Finally, in Chapter 4, we build on previous research highlighting boundary conditions for loss aversion, focusing on the magnitude of wins and losses. Previous research demonstrated that loss aversion may be reduced when the magnitude of outcomes is small, whereas the typical observation of loss aversion is present for large magnitude outcomes (Erev, Ert, & Yechiam, 2008; Harinck, van Dijk, Van Beest, & Mersmann, 2007). However, based on previous research demonstrating that people may be inaccurate in estimating their emotional reactions to future events (Gilbert, Morewedge, Risen, & Wilson, 2004; Wilson & Gilbert, 2003), we suggest that the effect of outcome magnitude on gain/loss asymmetries depends on whether participants are *predicting* how they will react to gains and losses versus their actual behavior in seeking gains and minimizing losses. In the work demonstrating that for small magnitude outcomes "gains loom larger than losses," participants were asked to rate the pleasantness/unpleasantness of losing or gaining different amounts of money and hypothetical gambles where participants indicate how much money they would be willing to risk in order to participate in a coin-toss gamble (e.g., Harinck et al. 2007). While in the present study, participants' decision-making actually led to experiencing gains and losses.

Limitations and Future Directions

The present investigation further demonstrates the robustness of the learning asymmetry. while also identifying conditions that modulate the magnitude of this asymmetry. Nonetheless, there are several limitations in the present investigation and questions that remain to be addressed in future research. First, we suggest that there are factors influencing loss learning that may be at play when the asymmetry is reduced. One possibility is that experience plays a role in reducing the asymmetry in that participants acquire some knowledge about the task structure that they incorporate to improve loss learning. However, what knowledge participants acquire remains unclear. Therefore, future studies can provide insight into what knowledge participants acquired from the first session. Another possibility is that with experience, participants adopt a different strategy. The finding that greater exploitation was associated with a smaller learning difference between wins and losses during the second session suggests that it is beneficial for participants to exploit what they learned rather than continue exploring different options. One approach to encourage this shift from exploring to exploiting is to introduce two phases in the

VLT. During the first phase, participants could be instructed to explore both images within each pair to learn the outcomes associated with each image. Then, during the second phase, participants could be instructed to use what they learned to select the “optimal” image. If this shift from exploring to exploiting is an ideal strategy, then the learning asymmetry should be reduced during the second phase.

Another factor that may be at play in reducing the learning asymmetry is the extent to which participants discount wins and losses. In Chapter 3, the learning asymmetry was eliminated when win and loss trials were separated but specifically, when win trials preceded loss trials. We suggest that without prior exposure to wins, participants may have discounted losses, thus, leading participants to focus on obtaining wins. On the other hand, when participants experienced wins before losses, participants may have focused on retaining their earnings, thus making losses become more meaningful. In Chapter 4, when varying the magnitude of outcomes, it is possible that the large magnitude payoffs selectively magnified large wins. Thus, participants discounted both losses and small wins. This may have resulted in small wins and losses to become equally meaningless due to the low magnitude. These findings suggest that participants may be valuing the wins more than losses. One approach to test this possibility is to find a loss value such that the loss will no longer be discounted. For example, participants may treat a loss of 10 points similarly to a win of 5 points. This possibility would coincide with research documenting that in order for losses and gains to be perceived similarly, losses should be at least twice the amount as wins (Kahneman & Tversky, 1979). Therefore, future work should examine what loss value would equate wins and losses in the VLT.

In addition, the question remains to what extent the present findings extend to other versions of probabilistic value learning tasks. For example, while the version of the VLT used in the present dissertation incorporates outcome probabilities of 20%/80%, other studies have used different probabilities, such as 33%/66% (Knutson, Samanez-Larkin, G. R., & Kuhnen 2011) and 40%/60% (Kim, Shimojo, & O'Doherty, 2006). Second, the number of win and loss pairs have varied. While our standard version of the VLT uses one win and one loss pair, other studies have used more than one pair (e.g., Raymond & O'Brien, 2009; Painter, Kritikos, and Raymond, 2014). While the learning advantage for win associations has also emerged in these studies, whether the context effects found in the present set of experiments included in this dissertation generalize to different variations of the VLT has yet to be tested.

While the results from the present dissertation provide evidence that the context in which win and loss associations are learned and the magnitude of win and loss outcomes can alter the learning pattern, it remains to be established whether these effects would play a role in how participants process the stimuli in subsequent encounters. That is, learning differences between win and loss associations could be the basis for valence effects rather than acquired value per se, however this remains to be tested. Thus, future work can examine how conditions that minimize the learning asymmetry affect responses to the same stimuli when they appear in subsequent cognitive and motor tasks.

Lastly, our goal of administering the memory task was to examine explicit knowledge of the outcomes associated with each VLT scene. However, the response options in the memory task rely on some assumptions about participants' interpretations of the probability of outcomes. For example, we designed "occasional win" to correspond to 20% win and "very likely to win" to 80% win. Yet, participants can differ in the extent to which they believe a win that occurs 80% of the time is "very likely" or "occasional." This variability could be one reason why performance on the memory task appears to suggest that people had poor memory of the probability of outcomes. The memory task could be refined to avoid such assumptions. One possible adjustment is rather than having categorical responses, participants can use a sliding scale that varies from -100% to 100% to indicate the probability that they believe the scene leads to a positive or negative outcome. Thus, participants would have to make a judgment about whether the scene leads to a win or loss and estimate the probability of the outcome. We could then compare the probabilities they select for the 20%/80% win and loss scenes to gauge their knowledge of the probability that the scene led to a win or loss.

Closing Remarks

The experiments in this dissertation were conducted to investigate boundary conditions for the learning advantage of win associations relative to loss associations. While we demonstrated that the learning asymmetry is remarkably robust, conditions were identified that reduced or eliminated the asymmetry. First, prior experience with the task reduced the asymmetry, at least quantitatively. Second, the context in which learning takes place affected the presence or absence of a learning asymmetry. Third, including multiple outcome magnitude conditions in the same context affects the efficiency of learning, as well as the learning pattern for win and loss associations. We also found that even when learning is equivalent, participants' explicit knowledge of the outcomes associated with wins and losses differed. Thus, equivalent endpoints do not necessarily mean equivalent learning. These findings suggest that in the context of probabilistic learning, people have a tendency to focus on win learning rather than loss learning. This may reflect an overall tendency to focus on positive outcomes while discounting negative outcomes. Future research can examine whether selectively increasing the loss value would equate the perception of wins and losses. Furthermore, research examining how learned value affects subsequent processing of stimuli previously associated with wins and losses should consider how choice history plays a role in subsequent effects of learned value.

References

Aberg, K. C., Doell, K. C., & Schwartz, S. (2016). Linking individual learning styles to approach-avoidance motivational traits and computational aspects of reinforcement learning. *Plos ONE*, *11*(11), e0166675.doi:10.1371/journal.pone.0166675

Anderson B. A. (2015). Value-Driven Attentional Capture is Modulated by Spatial Context. *Visual cognition, 23*(1-2), 67–81. https://doi.org/10.1080/13506285.2014.956851

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5,* 323–370. http://dx.doi.org/10.1037/1089-2680.5.4.323

Erev, I., Ert, E., & Yechiam, E. (2008). Loss aversion, diminishing sensitivity, and the effect of experience on repeated decisions. *Journal of Behavioral Decision Making, 21*(5), 575–597. https://doi.org/10.1002/bdm.602

Frank, M.J. & Hutchison, K. (2009). Genetic contributions to avoidance-based decisions: striatal D2 receptor polymorphisms. *Neuroscience, 164*(1), 131-140.

Frank, M.J., Seeberger, L.C., & O'Reilley, R.C. (2004) By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science, 306*(5703), 1940-1943.

Gilbert, D.T., Morewedge, C.K., Risen, J.L., & Wilson, T.D. (2004). Looking forward to looking backward: The misprediction of regret. *Psychological Science, 15*, 346–350.

Hao, C., Cabrera, L., Lin, Z., Reuter-Lorenz, P. A., & Lewis, R. L. (July 2019). Explaining Valence Asymmetries in Human Learning in a Symmetric Value Learning Paradigm: A Computational Reinforcement Learning Account. Presented at the 4th Multidisciplinary

Conference on Reinforcement Learning and Decision Making, Montréal, Québec, Canada.

Harinck, F., Van Dijk, E., Van Beest, I., & Mersmann, P. (2007). When gains loom larger than losses: Reversed loss aversion for small amounts of money. *Psychological Science*, *18*(12), 1099–1105. https://doi-org.proxy.lib.umich.edu/10.1111/j.1467-9280.2007.02031.x

Hochman, G., & Yechiam, E. (2011). Loss aversion in the eye and in the heart: The autonomic nervous system's responses to losses. *Journal of Behavioral Decision Making, 24*(2), 140–156. https://doi.org/10.1002/bdm.692

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263–291.

Kim, H., Shimojo, S., & O'Doherty, J. P. (2006). Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. *PLoS Biology, 4*(8), e233.

Knutson, B., Samanez-Larkin, G. R., & Kuhnen, C. M. (2011). Gain and loss learning differentially contribute to life financial outcomes. *PLoS One, 6*(9), e24390. https://doi.org/10.1371/journal.pone.0024390.

Ledgerwood, A., & Boydstun, A. E. (2014). Sticky prospects: loss frames are cognitively stickier than gain frames. *Journal of experimental psychology. General*, *143*(1), 376–385. https://doi.org/10.1037/a0032310

Lin, Z., Cabrera-Haro, L. E., & Reuter-Lorenz, P. A. (2020). Asymmetrical learning and memory for acquired gain versus loss associations. Cognition, *202*:104318. https://doi.org/10.1016/j.cognition.2020.104318

McGraw, A. P., Larsen, J. T., Kahneman, D., & Schkade, D. (2010). Comparing gains and losses. *Psychological science*, *21*(10), 1438–1445. https://doi.org/10.1177/0956797610381504

Nowak, M., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit- for-tat in the Prisoner's dilemma game. *Nature, 364*(6432), 56. https://doi.org/10. 1038/364056a0.

O'Brien, J. L., & Raymond, J. E. (2012). Learned predictiveness speeds visual processing. *Psychological Science, 23*(3), 359-363. doi: 10.1177/0956797611429800

Painter, D. R., Kritikos, A., & Raymond, J. E. (2014). Value learning modulates goal- directed actions. *The Quarterly Journal of Experimental Psychology, 67*(6), 1166–1175.

Raymond, J. E., & O'Brien, J. L. (2009). Selective Visual Attention and Motivation: The Consequences of Value Learning in an Attentional Blink Task. Psychological Science, 20(8), 981–988. https://doi.org/10.1111/j.1467-9280.2009.02391.x

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review, 5,* 296–320. http://dx.doi.org/10.1207/S15327957PSPR0504_2

Rutherford, H. J. V., O'Brien, J. L., & Raymond, J. E. (2010). Value associations of irrelevant stimuli modify rapid visual orienting. *Psychonomic Bulletin & Review*, *17*(4), 536–542. https://doi-org.proxy.lib.umich.edu/10.3758/PBR.17.4.536

Sparks, J., & Ledgerwood, A. (2017). When good is stickier than bad: Understanding gain/loss asymmetries in sequential framing effects. *Journal of experimental psychology. General*, *146*(8), 1086–1105. https://doi.org/10.1037/xge0000311

Theeuwes, J., & Belopolsky, A. V. (2012). Reward grabs the eye: oculomotor capture by rewarding stimuli. *Vision research, 74*, 80–85. https://doi.org/10.1016/j.visres.2012.07.024

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science (New York, N.Y.)*, *315*(5811), 515–518. https://doi.org/10.1126/science.1134239

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin, 134*(3), 383-403. doi: 10.1037/0033-2909.134.3.383

Wilson, T. D., & Gilbert, D. T. (2003). Affective forecasting. *Advances in Experimental Social Psychology, 35*, 345– 411. http://dx.doi.org/10.1016/S0065-2601(03)01006-2

Yechiam, E., & Hochman, G. (2013). Losses as modulators of attention: Review and analysis of the unique effects of losses over gains. *Psychological Bulletin, 139*(2), 497–518. https://doi.org/10.1037/a0029383